

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 11:14:25 ; Search time 6119 Seconds
(without alignments)
10783.989 Million cell updates/sec

Title: US-09-647-841B-1
Perfect score: 1613
Sequence: 1 ttttataatatttaagctt.....tgaacttgagtaccttactc 1613

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.in.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|---------------------|
| 1 | 1428.6 | 88.6 | 3617 | 8 | AF028841 | X6028841 Medicago |
| 2 | 175.4 | 10.9 | 3234 | 8 | PSPTXAG | X67427 Pisum sativa |
| 3 | 117.8 | 7.3 | 19520 | 2 | BX088600 | DX088600 Danio rer |
| 4 | 115.6 | 7.2 | 128769 | 2 | AL139261 | AL139261 Homo sapi |
| c 5 | 114.6 | 7.1 | 129240 | 9 | AC084128 | AC084128 Homo sapi |
| c 6 | 114 | 7.1 | 8056 | 6 | AX599046 | AX599046 Sequence |
| c 7 | 113.2 | 7.0 | 123589 | 9 | AC104790 | AC104790 Homo sapi |
| c 8 | 113.2 | 7.0 | 172758 | 2 | AC022553 | AC022553 Homo sapi |
| c 9 | 112.8 | 7.0 | 159948 | 2 | BX248240 | BX248240 Danio rer |
| 10 | 112.8 | 7.0 | 187310 | 2 | BX005122 | BX005122 Danio rer |
| c 11 | 112.2 | 7.0 | 178066 | 2 | BX004966 | BX004966 Danio rer |
| 12 | 111.2 | 6.9 | 112917 | 8 | AC123976 | AC123976 Medicago |
| c 13 | 110.8 | 6.9 | 196490 | 2 | AC005507 | AC005507 Plasmodiu |
| c 14 | 110.6 | 6.9 | 147956 | 2 | AC137839 | AC137839 Medicago |
| 15 | 109.8 | 6.8 | 195032 | 2 | AC144827 | AC144827 Danio rer |
| 16 | 109 | 6.8 | 169546 | 2 | AC004157 | AC004157 Plasmodiu |
| c 17 | 109 | 6.8 | 250707 | 3 | AE014848 | AE014848 Plasmodiu |
| c 18 | 108.8 | 6.7 | 180919 | 2 | BX001038 | BX001038 Danio rer |
| c 19 | 108.2 | 6.7 | 196859 | 9 | AC087277 | AC087277 Homo sapi |
| c 20 | 107.4 | 6.7 | 80150 | 9 | AC134620 | AC134620 Homo sapi |
| c 21 | 107.2 | 6.6 | 26561 | 9 | AL929515 | AL929515 Human DNA |
| c 22 | 106.8 | 6.6 | 181792 | 9 | AC098822 | AC098822 Homo sapi |
| 23 | 106.6 | 6.6 | 136713 | 9 | AC112255 | AC112255 Homo sapi |
| c 24 | 106.4 | 6.6 | 241454 | 2 | BX470172 | BX470172 Danio rer |
| 25 | 105.8 | 6.6 | 151815 | 2 | BX470147 | BX470147 Danio rer |
| 26 | 105.6 | 6.5 | 101884 | 8 | AC142224 | AC142224 Medicago |
| c 27 | 105.6 | 6.5 | 188485 | 2 | BX470246 | BX470246 Danio rer |
| c 28 | 105.6 | 6.5 | 188716 | 5 | AL807739 | AL807739 zebrafish |
| 29 | 105.4 | 6.5 | 198583 | 2 | AC144702 | AC144702 Danio rer |
| c 30 | 105.2 | 6.5 | 60529 | 2 | AC144825 | AC144825 Danio rer |
| c 31 | 104.8 | 6.5 | 89728 | 5 | AL672079 | AL672079 zebrafish |
| 32 | 104.8 | 6.5 | 126175 | 2 | AC136841 | AC136841 Medicago |
| 33 | 104.6 | 6.5 | 201591 | 2 | AL954657 | AL954657 Danio rer |
| 34 | 104.6 | 6.5 | 313050 | 3 | PFA929352 | AL929352 Plasmodiu |
| c 35 | 104.2 | 6.5 | 187197 | 2 | BX530723 | BX530723 Danio rer |
| 36 | 104.2 | 6.5 | 129240 | 9 | AC084128 | AC084128 Homo sapi |
| c 37 | 104 | 6.4 | 343050 | 3 | PFA929353 | AL929353 Plasmodiu |
| c 38 | 103.6 | 6.4 | 67970 | 3 | PFMAL1P3 | AL031746 Plasmodiu |
| c 39 | 103.6 | 6.4 | 250078 | 3 | AE014829 | AE014829 Plasmodiu |
| 40 | 103.4 | 6.4 | 1867 | 8 | MTSCAJ23 | AJ223323 Saccharom |
| c 41 | 103.4 | 6.4 | 2144 | 5 | BC044278 | BC044278 Xenopus l |
| c 42 | 103.4 | 6.4 | 146039 | 5 | AL627132 | AL627132 zebrafish |
| c 43 | 103.2 | 6.4 | 130110 | 9 | AL731777 | AL731777 Human DNA |
| 44 | 103.2 | 6.4 | 152369 | 2 | CNS07ED2 | AL445443 Homo sapi |
| 45 | 103.2 | 6.4 | 172816 | 9 | AC093899 | AC093899 Homo sapi |

ALIGNMENTS

RESULT 1
AF028841
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AF028841
Medicago sativa
AF028841.1 GI:3818415
Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1224 to 3617)
Deutsch,C.E. and Winicov,I.

Pred. No. is the number of results predicted by chance to have a

[illegible]


```

Db 32729 TATAAAATATAT-----ATATAAATATATATAAATATATATAAATATATATATAAATTTATATATATAG 32674
QY 305 CGATTCTTCTTGTCGACGCTTGGAGAACCCCTATCTGGCGCTTGGAGAGATTTACTTCTGTG 364
Db 32673 ATATATAAATAGATATATATTTATATATAAATATATTTATATATAAATATATAAATATATAT 32614
QY 365 TGATGCTTCTAGACGACGCTTGAAGGCTGTAGTCTAGTT---TTTTTTTTCATCCTT 421
Db 32613 TAATATATATAAATATATAAATTTATATATAAATTTATATATAAATATATAAATATATCAATGTAT 32554
QY 422 CCTACCAAAAAAAGAGTCATATAATAGCTTTATACATATACTTTATACATATACTTTAATAAATAA 481
Db 32553 AATATATTATATATATATTTATATATATATTTTATATATAAATATATAAATATATAAATATA 32494
QY 482 AAAATTTCATCCCTAAACACAGTAGAGAAATTTTCATRAAAAAAATATTTGTTTATAAATTA 541
Db 32493 TATA-TAAATATATATAAATATATAAATTTATATATAAATATATAAATTTATATATATAT 32435
QY 542 CATGCCGTTACGGTAAATAATGGATAATTTGGGTATGGAGTACGTAAATTAAGGTT 601
Db 32434 AAATATATAAATTTATATATAAATATATAAATTTATATATATGTTTATATATATAAATATAT 32375
QY 602 CATTTGGTTAAAAAAGTAAATAAATTTCTCTCCTGATTTATATGAAATGACATTTTTT 661
Db 32374 AATATATATAATTTATATATAAATATAT-----ATAAATTTATATAAATTTATATAATTT 32320
QY 662 TGGACATGAAGGCTATTGATTTTACCACCTTTTACACCTTTTCAAGCCATTCGAAGAT 721
Db 32319 ATATATATATAATATATATATTTATATATAATACATATAAATAAATTAATATATAT 32260
QY 722 GAATATAGATTTTGGGCGATCAACACAGAAGATCATTACGATAACATGCTTTATACATAC 781
Db 32259 AAATATATATTTATAAATAAATATATAAATTAATAAATATATATTTATAAATAAATATAT 32200
QY 782 CCCGTCATCTCTTTTACCACATAAATGAAATGCTTCTGCTTTTCGTTAAGCA 841
Db 32199 ATAAATATAAATATATATATAAATAAATATAAATATAAATAAATA-TATATTATAA 32141
QY 842 TAAAAACATCAAGCTAGCAAAAATGTTGTTTGGCGATGACACATTCATATAGTTTAA 901
Db 32140 TAAATATATAAATATAAATATATATTTATAAATAAATATATAAATATAAATATAAATA 32081
QY 902 AGGATGATGATTCGATTACAAAAACAAATPACTAATAATCTAGCACAAAGTTTAAAGC 961
Db 32080 TTTATAAATAAATATTTATAAATAAATATAAATAAATAAATAAATAAATAAATAAATA 32021
QY 962 AATATTATAAGCTTCATAGCATCTGGATATTCATTTAGAAATATAGAT 1010
Db 32020 TATATATTATAAATAATATGATTTATATAAATAAATAAATAAATAAATAAATAAATA 31972

```

```

RESULT 6
AX599046/C. AX599046 8056 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 386 from Patent WO02077272.
DEFINITION AX599046
ACCESSION AX599046
VERSION AX599046.1 GI:28399186
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Berlin,K., Braun,A., Distler,J., Guetli,D., Howe,A., Mueller,J.,
Olek,A., Piepenbrock,C., Adorjan,P., Grabs,G., Lesche,R., Leu,E.,
Lewin,A., Lipscher,E., Maier,S., Model,F., Mueller,V., Otto,I.,
Pelet,C. and Ziebarth,H.
TITLE Methods and nucleic acids for the analysis of hematopoietic cell
proliferative disorders
JOURNAL Patent: WO 02077272-A 386 03-OCT-2002;
EpiGenomics AG (DE)
FEATURES
source Location/Qualifiers
1..8056
/organism="synthetic construct"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 3711 a 0 c 371 g 3974 t
ORIGIN
Query Match 7.1%; Score 114; DB 6; Length 8056;
Best Local Similarity 44.6%; Pred. No. 2.9e-06;
Matches 613; Conservative 1; Mismatches 751; Indels 9; Gaps 4;
QY 1 TTTTATAAATATATTTAAGCTTGCATATAATTTTGGCATCTATATATAAGCCCACTACCAAT 60
Db 1855 TTTTATTTTATTTTATTTTAAACAAAATTTATTTATTTATTAATTAATAATAATA 1796
QY 61 TTAATAATATATATATATATATATATATATATATATATATATATATATTTTATTTACCAAT 120
Db 1795 ACATAATTAATTTTAAATTTTAAATTTTAAATAATTTTATTTTAAATTTTAAATAA 1736
QY 121 TTAATAATATATATATATATATATATATATATATATATATATATATATTTTATTTATTTA 180
Db 1735 TTAATAATTAATAATATATATATATATATATATATATATATATATTTTATTAATAA---A 1679
QY 181 TTAGGTCGTCGTAAGGTAAGTAATTAATTAATTTTGTACCATTTAAAGTCATATAATAGT 240
Db 1678 TAAATTTTATCATATATTAATTAATAATAATTTTATTTTAAATTTTACAAATTCATTTATTTCAA 1619
QY 241 ACAATCCACCCCTTTGAGAGGTTAATGTGTGCGGATTTTCTAGATAAACAAGCGCCA 300
Db 1618 TTTTAAATATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATTTT 1559
QY 301 TTCAGATCTCTTGTGTCAGCTTGGAGAACCCCTATCTCGGCTTGGAGATTTTACATTC 360
Db 1558 TAAAAAATAATTTTATTTAAATAATTTTAAACATATTTATATTTATATATATAT 1499
QY 361 TTGTTGATGCTCTAGAGTACAGCTCCTTAGGCTGTAGTCTAGTCTTTTCTCATCCT 420
Db 1498 ATTAATTTTAATAATTAATTTTAAATAATTTTATTTAAATAAATAAATAATTTTATTTT 1439
QY 421 TCCTACCAAAAAAAGTCAATAATATAGTTTATACATATACTTTATATAAATAAATAA 480
Db 1438 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1379
QY 481 AAAAATTTTCATCCCTAAACATAGTAGAAAATTTTCATAAAAAATAATTTGTTTATAATTT 540
Db 1378 AATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1319
QY 541 ACATGCCGTTACGGTAAATAATGGATAATTTGGGTATGGAGTACTAGTAATTAATAAGT 600
Db 1318 TTTTATATATTTTATTAATTTTAAAAAATAATTTACATTTTATATATATATATATATAT 1259
QY 601 TCATTGGTTAAAAAACCTAAAAATAATTTCTCTCGATTTATATGAAATGACATTTT 660
Db 1258 TAACATTTTATTTTCAAAAAAATAATTTTATTTTAAATAAATAAATAAATAAATAA 1199
QY 661 TTGGAACATGAAGGCTATTTTACACCTTTTACACCTTTTCAAAAGCCATTTCAAGGA 720
Db 1198 TTAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1139
QY 721 TGAATATAGATTTTGGCGGATCAACACAGAATCATTTAGATAACATGCTTATACATA 780
Db 1138 TCAAAAAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1079
QY 781 CCCGTCGATCTCTTTTATCCCAATAACATTTGAAATGTTGCTTCTTTTCGTTAAGC 840
Db 1078 TAAACAAATAATTTTATTTTATTTTACATAATTAATAATTTCAAAATTAATAAATAAATAA 1019
QY 841 ATAAAAACATCAAAAGTCTAGCAAAATGTTGTTTGGCATGACATTCATATATAGTTTA 900
Db 1018 TTTTAAATAAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 963
QY 901 AAGGATGATGATTCGATTAACAAACAAATACTATAATTTCTAGCAAAAGTTTAAAG 960
Db 962 AAAAAATAAATAATTTTATTTTAAAAAATAATTTTATTAACAAAAACAACAAAAATAAATAA 903

```

```

QY 961 CAATATTAAAGCTTCATAGCATGGATGATTCATTAGAAAATATAGATTGCCC 1020
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 TATAAATTTAAATTAATACAAATTTTATACATAAAAAAATAAAAAACAATTTT 843
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 CTTTCATCAGCGGTCTAACAGCACCACTGCTCACTACATGTCAAAAATGCTCTAGTAC 1080
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 TTTAAATAAAAATATTTTAAATAATCATTTTATAAATTTTATTTAAATATAAATAAAT 783
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 AGCACCGCTTTTACTTGATTCCTCCCTGTCATGCAAGAAATCAAAACAATATTG 1140
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 AATTTTATAATTTAATTAATTTTAAATTTAAAAATTTAATTAATAAATAAATTTT 723
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 GACACACAACTTGCCTCCCACTTCTCTTTCTGCTGCTAGTTGTTGAGACTCAT 1200
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 TAAATTAATTTTAAATAAATAAATAAATTTTATATATTTATTTTATTTTATTTT 663
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1201 ATTGA-TCAATTTGCTGATGAATTCACACAAAAATTCACCTACCCATTCGATGCTG 1259
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATAAATTTTATTTAAAAATTTATTTTAAAAATTTTAAAAATTTTATATATTTT 603
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1260 GGGCCACATATAAATCCATGAAGGATTCAT-GTCCATCCCAAGTCAATGATTCACAT 1318
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 TAAATATTATTAATAATAAATTTTAAATAAATTTTAAATAAATTTATATAAATAAAT 543
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1319 ATATAACATGAATAATTTAATTCCAATTTGCAGTATTATGATTTAGATTGATT 1372
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 TTTAAATTTAATAATAAATTTAATTAATTAATAAATTTAATAAATTTTAAATTTT 489
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AC104790/c
LOCUS Homo sapiens BAC clone RP11-167A8 from 4, complete sequence.
DEFINITION AC104790
ACCESSION AC104790.4 GI:20258589
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 123589)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 123589)
Pearman,C., Meyer,R. and Doeber,A.
The sequence of Homo sapiens BAC clone RP11-167A8
Unpublished (2001)
REFERENCE 3 (bases 1 to 123589)
Waterston,R.H.
Direct Submission
Submitted (21-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 123589)
Waterston,R.H.
Direct Submission
Submitted (21-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 123589)
Waterston,R.
Direct Submission
Submitted (29-MAY-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 21, 2002 this sequence version replaced gi:18652549.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc

```

Contact: sapiens@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0167A08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Cataneese, J.J. and de Jong, P.J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-460L16; the clone sequenced to the right is RP11-162F1, 2000 bp overlap. Actual start of this clone is at base position 41945 of RP11-460L16; actual end is at base position 44749 of RP11-162F1.

Polymorphisms exist between AC104790 and AC104804. Data from AC104804 and AC106867 was used to finish AC104790.

FEATURES

| | source |
|---------------|--------------------------|
| repeat_region | 1. .123589 |
| | /organism="Homo sapiens" |
| | /mol_type="genomic DNA" |
| | /db_xref="taxon:9606" |
| | /chromosome="4" |
| | /map="4" |
| | /clone="RP11-167A8" |
| | /clone_lib="RPCT-11" |
| | 353. .379 |
| repeat_region | /rpt_family="AT-rich" |
| | 1639. .1719 |
| repeat_region | /rpt_family="L1" |
| | 1760. .1953 |
| repeat_region | /rpt_family="L1" |
| | 1992. .2398 |
| repeat_region | /rpt_family="L1" |
| | 4331. .4798 |
| repeat_region | /rpt_family="ERV1" |
| | 5422. .5963 |
| repeat_region | /rpt_family="ERV1" |
| | 5964. .6006 |
| repeat_region | /rpt_family="(TG)n" |
| | 6007. .6703 |
| repeat_region | /rpt_family="ERV1" |
| | 7219. .7695 |
| repeat_region | /rpt_family="ERV1" |
| | 7860. .7909 |
| repeat_region | /rpt_family="L2" |
| | 7910. .8380 |
| repeat_region | /rpt_family="MALR" |

[illegible]

RESULT 10

| | | | | | |
|----------|------------|---|-----|--------|-----------------|
| EX005122 | LOCUS | 187310 bp | DNA | linear | HTG 02-JUN-2003 |
| | DEFINITION | Danio rerio clone CH211-134D11, *** SEQUENCING IN PROGRESS ***, 4 unordered pieces. | | | |

ACCESSION BX005122

VERSION BX005122.7 GI:29500520

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii: Teleostei: Ostariophysi:

Cypriniformes: Cyprinidae: Danio.

REFERENCE 1 (bases 1 to 187310)

AUTHORS McLaren, S.

| TITLE | Direct Submission |
|-----------|-------------------|
| THEME | |
| NUMBER/01 | |

JOURNAL
Submitted (29-MAY-2003) Wellcome Trust Sanger Institute, Hinxton

Cambridgeshire CB10 1SA UK E-mail enquiries:
 Cambridge (25 May 2005) welcome trust sanger institute, Hinxton,

Cambridge, CB2 3RQ, UK. E-mail enquiries:
zfish-help@cam.ac.uk Clone requests: clonerequests@cam.ac.uk

COMMENT

COMMENT
On Apr 2, 2003 CILTS sequence version replaced g1:28272911.
----- Genome Center

Center: Wellcome Trust Sanger Institute

Center: Wellcome Trust Sanger Institute
Center code: 00

Center code: SC
Web site: <http://www.southcentral.edu>

web site: <http://www.sanger.ac.uk>

Contact: zf1sh-help@esanger.ac.uk

----- Project Information

Center project name: zC134D11

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 184982 bases at least Q40

Consensus quality: 185170 bases at least Q30

Consensus quality: 185407 bases at least Q20

Insert size: 187010; sum-of-contigs

Insert size: 190203; 4.5% error; aqarose-fp

Quality coverage: 9.35x in 020 bases; sum-of-contigs Quality

coverage: 9.31x in 020 bases: agarose-fp

| Year | 1980 | 1981 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1980 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 | | | | | | | | | | | | | | | | | | | | | |

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

- * is not known and their order in this sequence record is
- * arbitrary. Gaps between the contigs are represented as
- * runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence.
- * as soon as it is available and the accession number will
- * be preserved.

| | | | |
|---|--------|---------|-------------------------------|
| * | 1 | 98152: | contig of 98162 bp in length |
| * | 98163 | 98562: | gap of 100 bp |
| * | 98263 | 135335: | contig of 36973 bp in length |
| * | 132326 | 135335: | gap of 100 bp |
| * | 135336 | 140884: | contig of 5549 bp in length |
| * | 140885 | 149984: | gap of 100 bp |
| * | 140985 | 187310: | contig of 46326 bp in length. |

[illegible]


```

REFERENCE
AUTHORS      2 (bases 1 to 112917)
Shaul,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (06-JUN-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS      3 (bases 1 to 112917)
Shaul,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (12-NOV-2002) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS      4 (bases 1 to 112917)
Shaul,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (08-WAR-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS      5 (bases 1 to 112917)
Shaul,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (11-JUN-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS      6 (bases 1 to 112917)
Shaul,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (11-JUN-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT      On Jun 11, 2003 this sequence version replaced gi:31560173.
-----
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
1..112917
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-30123"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 38546 a 17840 c 18400 g 38131 t
ORIGIN
Query Match 6.9%; Score 111.2; DB 8; Length 112917;
Best Local Similarity 68.5%; Pred. No. 3.3e-06;
Matches 152; Conservative 1; Mismatches 69; Indels 0; Gaps 0;
QY 235 TATAGTACATCAACCCCTTTGAGAGGTTAATGTGTGCGGATTTCTAGATAACAAG 294
Db 22909 TGTATGTGACACACACTCTTGTGAGAGGTTAAACAAGTGTGCAGATTCCTTAGTACGCTTA 22968
QY 295 GYCCATTCACGATTCCTTGTGCGAGCTTGAGAACCCCTATCCTGGCGTGGAGATT 354
Db 22969 GTGCTATTGATGATTCCTTGTGCGAGCTTGTGAACTCATCTTATGCTTGGAGATT 23028
QY 355 TACTTCTTGTGATGCTTCTAGAGTACAGCTCCTTAAAGGCTGTAGTCTAGTCTATTTT 414
Db 23029 TACTTATTGATGCTTCTAGAGTGCAGTCTTCTAGGTTATTAACCTAGTCTTTCTTCT 23088
QY 415 CATCCTTCTACAAAAAAGTCAFAAATATAGTTTA 456
Db 23089 TCCTGCACTAGTAATAAAGATAAATAGTAACAAAGATTA 23130
RESULT 13

AC005507/c
LOCUS
DEFINITION
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
AC005507
VERSION
AC005507.8 GI:9797737
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 196490)
Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kurd,I.O.B., Conway,A.B. and Davis,R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 196490)
Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
Direct Submission
TITLE
Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT
On Aug 12, 2000 this sequence version replaced gi:8810456.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 181303: contig of 181303 bp in length
* 181304 181503: gap of unknown length
* 181504 196490: contig of 14987 bp in length.
* Location/Qualifiers
1..196490
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFYACB8-628"
/clone="3D7"
BASE COUNT 77103 a 20186 c 18680 g 80315 t 206 others
ORIGIN
Query Match 6.9%; Score 110.8; DB 2; Length 196490;
Best Local Similarity 64.5%; Pred. No. 3.2e-06;
Matches 182; Conservative 0; Mismatches 97; Indels 3; Gaps 1;
QY 2 TTTATAAATATTTAAGCTTGTGATAAATAATTTTGGGATCTATATATA---AGCCCACTACCA 58
Db 182646 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182587
QY 59 ATTATAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 118
Db 182586 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182527
QY 119 ATTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 178
Db 182526 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 182467
QY 179 TATTACGTTGATGGTAAAAAATAAATAAATTTGTTACCTTTAAAGTCATAAATATA 238
Db 182466 ATTTATAAGAAATGTACATAAACAATAACGTTCAATTAATTTTAAATCATTTATAA 182407
QY 239 GTACAATCAACCCCTTTGAGAGGTTAAATGCTGTGCGGATTT 280
Db 182406 ATATCTTACCTTCATGTTAATTTTAAATGTTTAAAGAATTT 182365

RESULT 14
AC137839/c
LOCUS
DEFINITION
Medicago truncatula clone mch2-34j11, WORKING DRAFT SEQUENCE, 18

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

unordered pieces.
AC137839
AC137839.10 GI:31076555
HTG; HTGS_PHASE1; HTGS_DRAFT.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-34j11
Unpublished
2 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (04-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 147956)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (27-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 27, 2003 this sequence version replaced gi:30270629.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

COMMENT

NOTE: This is a 'working draft' sequence. It currently
consists of 18 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 2190: contig of 2190 bp in length
2191 2290: gap of unknown length
2291 5232: contig of 2942 bp in length
5233 5332: gap of unknown length
5333 7396: contig of 2064 bp in length
7397 7496: gap of unknown length
7497 10640: contig of 3144 bp in length
10641 13741: gap of unknown length
13741 13777: contig of 2637 bp in length
13778 13478: gap of unknown length
13478 16887: contig of 3210 bp in length
16888 16787: gap of unknown length
16788 19875: contig of 3088 bp in length
19876 19975: gap of unknown length
19976 23802: contig of 3827 bp in length
23803 23902: gap of unknown length
23903 27548: contig of 3646 bp in length
27549 27648: gap of unknown length
27649 33103: contig of 5455 bp in length
33104 33203: gap of unknown length
33204 41209: contig of 8006 bp in length
41210 41309: gap of unknown length
41310 47453: contig of 6144 bp in length
47454 47553: gap of unknown length
47554 57436: contig of 9883 bp in length
57437 57536: gap of unknown length
57537 67037: contig of 9501 bp in length
67038 67137: gap of unknown length
67138 78583: contig of 11446 bp in length
78584 78683: gap of unknown length

* 78684 95337: contig of 16654 bp in length
* 95338 95437: gap of unknown length
* 95438 121254: contig of 25817 bp in length
* 121255 121354: gap of unknown length
* 121355 147956: contig of 26602 bp in length.

FEATURES
source

Location/Qualifiers
1..147956
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-34j11"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 48249 a 24438 c 25475 g 48056 t 1738 others
ORIGIN

Query Match 6.9%; Score 110.6; DB 2; Length 147956;
Best Local Similarity 61.2%; Pred. No. 3.6e-06;
Matches 194; Conservative 1; Mismatches 120; Indels 2; Gaps 1;

QY 228 TCATAAATATAGTACAATCCAAACCTTTGAGAGGTTAATGTGTGCGGATTTCTAGAT 287
DB 111166 TGAATGTTCATTTTACCCACACCATTCGAGAGACAAATGAGTGTAGACTTTTAGCT 111107
QY 288 AAACAAGGCGCATTCACGATTCCTTTGGTGCAGCTGGAGAACCCATATCCTGGCTTG 347
DB 111106 AAGCATGTGCGCAATCAGACTCTTTCATGTCAGCTGGAGAACCCCTCTCCGGCTTG 111047
QY 348 GAAGATTTTACITCTTGTGTGATGCTTCTAGAGTACAGCTCTTAAAGCTGTAGCTAGTTT 407
DB 111046 GAAGCTTTTACTTCTCACAATGATGCTAGAGTGTAGTTTC--TTAGCGGCTAGCTAAATTT 110989
QY 408 TTTTTCATCTCTCTACCAAAAAAAGTATATAATAGTTTATACATATAACT 467
DB 110988 TCTGTTTTCCTTCTCTTCTATTTACCAAAAAACAAAAAGATATATAA 110929
QY 468 TTAATAAATAAATAAATAATTTATCCCTAAAAACATAGTAGATAAATTTTCAAAAAAATA 527
DB 110928 TCAGTTTCTAAAAAATAATACAGCATGMAAACTCTGAAAAAACCAAAACACATTTTC 110869
QY 528 TTGTTTATAATTACAT 544
DB 110868 CTATCTCAATGGATAT 110852

RESULT 15

AC144827
LOCUS AC144827 195032 bp DNA linear HTG 22-MAY-2003
DEFINITION Danio rerio clone CH211-77C11, *** SEQUENCING IN PROGRESS ***, 7
ordered pieces.
AC144827
ACCESSION AC144827.1 GI:30984602
VERSION HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
KEYWORDS Danio rerio (zebrafish)
SOURCE Danio rerio

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 195032)
Talbot,W.S., Rauch,G.J., Grimwood,J., Dickson,M., Schmutz,J. and
Myers,R.M.
Genomic studies of vertebrate diversity

JOURNAL

2 (bases 1 to 195032)
Talbot,W.S., Rauch,G.J., Grimwood,J., Dickson,M., Schmutz,J. and
Myers,R.M.
Direct Submission

TITLE

Submitted (22-MAY-2003) Stanford Human Genome Center, 975
California Avenue, Palo Alto, CA 94304, USA

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have

* provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 61516: contig of 61516 bp in length
 61616: gap of unknown length
 117922: contig of 56306 bp in length
 117923: gap of unknown length
 125117: contig of 7095 bp in length
 125217: gap of unknown length
 126859: contig of 1642 bp in length
 126959: gap of unknown length
 134961: contig of 8002 bp in length
 135061: gap of unknown length
 135062: contig of 12492 bp in length
 147554: gap of unknown length
 147654: 195032: contig of 47379 bp in length.
 The sequence of the clone was established as a mapping and
 sequencing collaboration at the Stanford Genome Evolution Center,
 funded by the NIH Centers of Excellence in Genomic Science (CEGS),
 initiative (http://cegs.stanford.edu). The clone was isolated from
 the BAC library CHORI-211 (http://bacpac.chori.org).

-----Genome Center
 Center: Stanford Human Genome Center
 Center Code: shgc
 Web site: http://www.shgc.stanford.edu

 Project Information
 Center Project Name: 2341
 Center clone name: CH211-77C11

 Summary Statistics
 Consensus quality: 191969 bases at least Q40
 Consensus quality: 192317 bases at least Q30
 Consensus quality: 192756 bases at least Q20
 Estimated insert size: 190000; agarose-fp estimation
 Estimated insert size: 193122; sum-of-contigs

estimation

Quality coverage: 1.01 in Q20 bases; agarose-fp

estimation
 Quality coverage: 1 in Q20 bases; sum-of-contigs

estimation.

FEATURES
 source

Location/Qualifiers
 1. 195032
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-77C11"

BASE COUNT 61347 a 35379 c 36159 g 61547 t 600 others
 ORIGIN

Query Match 6.8%; Score 109.8; DB 2; Length 195032;
 Best Local Similarity 42.9%; Pred. No. 4.3e-06;
 Matches 412; Conservative 0; Mismatches 542; Indels 7; Gaps 3;
 Qy 7 AATATTAAAGCTGTAATAAATTTGCGATCTATATAAGCCACTACCAATTTAAAA 66
 Db 134453 AAACAATAATGTGCACATTTAAATGAATTTCTATAGTTAATCATTCACAAATATATAT 134512
 Qy 67 TTAT 126
 Db 134513 AT 134572
 Qy 127 TTAT 185
 Db 134573 AT 134632
 Qy 186 TTGATGCTAAAAAATAAATAAATTTGTTTACCATTAAAAAGTCATAAATAATAGTACAAT 245
 Db 134633 AT 134692
 Qy 246 CCACCCCTTTGAGAGGTTAATGTGTGGCGGATTTTCTAGATAAACAAGGCGCATTCAC 305

Db 134693 AT 134752
 Qy 306 GATTCCTTTGGTGCAGCTTGGAGAACCCCTATCTCGGCTTGGAGATTTTACTTCTGTT 365
 -Db 134753 AT 134812
 Qy 366 GATGCTTCTAGAGTACAGCTCCTTAAGGCTGTAGTCTAGTCTAGTCTAGTCTAGTCT 425
 Db 134813 TATATTAT 134871
 Qy 426 CCAAAAAAAGAGTCATAAATATAGTTTATACATATATATATATATATATATATATATAT 485
 Db 134872 TAT 134931
 Qy 486 TTTTCATCCCTAAACACATAGTAAATTTTCAAAAAAATAATTTGTTTATATATATATAT 545
 Db 134932 TAAAAATTAAAAATTTTAT 134991
 Qy 546 CCGTTACGGTAAAAAATGGATAAATTTGGGTATGGAGTACTAGTAAATTAAGGTTTCATT 605
 Db 134992 NNN 135051
 Qy 606 GGTAAAAAACAATAAATAATTTCTCTCTGATTTATATGAATGACATTTTGGGA 665
 Db 135052 NNN 135111
 Qy 666 ACATGAAGGCTATTGATTTTACACCTTTTACACCTTTTCAAGCCATTCAGGATCAAT 725
 Db 135112 ATACATTTTACATAATATCTTTTTTTTATATATATATATATATATATATATATAT 135171
 Qy 726 ATAGATTTTGGGGGATCAACACAGAAATCATTCAGTAAATGCTTTTACATATATATATAT 785
 Db 135172 ATAAAT 135226
 Qy 786 TCAATCTCTTTTATTTTACCAATAAACATTTGAAATGCTGCTCTTTCGTTAAGCATAAA 845
 Db 135227 AAATTTTAT 135286
 Qy 846 AACATCAAGCTAGCAAAATGTTTTCGATGACACATTTTCATATATATATATATATATAT 905
 Db 135287 TCATATATAAT 135346
 Qy 906 TGCATGATTCGATTAACAAACAAATACTAATAATTTCTAGCACAAGTTTAAAGCAATA 965
 Db 135347 TAT 135406
 Qy 966 T 966
 Db 135407 T 135407

Search completed: August 14, 2003, 20:46:16
 Job time : 6137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 10:52:01 ; Search time 482 Seconds
(without alignments)
9033.597 Million cell updates/sec

Title: US-09-647-841B-1

Perfect score: 1613

Sequence: 1 ttttataaatatttaagctt.....tgaacttgagtaaccttactc 1613

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

| | |
|-----|--|
| 1: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.* |
| 2: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.* |
| 3: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.* |
| 4: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.* |
| 5: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.* |
| 6: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.* |
| 7: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.* |
| 8: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.* |
| 9: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.* |
| 10: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.* |
| 11: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.* |
| 12: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.* |
| 13: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.* |
| 14: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.* |
| 15: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.* |
| 16: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.* |
| 17: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.* |
| 18: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.* |
| 19: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.* |
| 20: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.* |
| 21: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.* |
| 22: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.* |
| 23: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.* |
| 24: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.* |
| 25: | /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 1427 | 88.5 | 1612 | 21 | AAZ34539 |
| c | 114 | 7.1 | 8056 | 25 | ABZ10246 |
| c | 100.6 | 6.2 | 16258 | 24 | ABL70376 |
| c | 100.6 | 6.2 | 16258 | 24 | ABK40038 |
| 5 | 100.2 | 6.2 | 8056 | 25 | ABZ10246 |
| c | 95.2 | 5.9 | 8056 | 25 | ABZ10100 |
| 7 | 93.2 | 5.8 | 6641 | 24 | ABL54336 |
| 8 | 93.2 | 5.8 | 6641 | 24 | ABL32315 |

| | | | | | | | |
|----|------|------|-----|--------|----|----------|--------------------|
| c | 9 | 92.4 | 5.7 | 9770 | 24 | ABL32032 | Human immune syste |
| 10 | 92 | 92 | 5.7 | 6621 | 24 | ABL70156 | Chemically treated |
| 11 | 92 | 92 | 5.7 | 6621 | 24 | ABK33933 | Human DNA for stag |
| 12 | 92 | 92 | 5.7 | 6621 | 24 | AAS61098 | Human gene regulat |
| 13 | 92 | 92 | 5.7 | 6621 | 25 | ABZ09984 | Haematopoietic cel |
| 14 | 92 | 92 | 5.7 | 6621 | 25 | ABZ10098 | Haematopoietic cel |
| 15 | 92 | 92 | 5.7 | 6621 | 25 | ABZ10130 | Haematopoietic cel |
| 16 | 92 | 92 | 5.7 | 6621 | 25 | ABZ10244 | Haematopoietic cel |
| 17 | 91.6 | 91.6 | 5.7 | 11745 | 24 | ABK28332 | DNA transcription |
| c | 18 | 91.4 | 5.7 | 6641 | 24 | ABL54336 | Chemically treated |
| c | 19 | 91.4 | 5.7 | 6641 | 24 | ABL32315 | Human immune syste |
| c | 20 | 91 | 5.6 | 3991 | 22 | AAD16633 | Human novel protei |
| c | 21 | 90.6 | 5.6 | 8056 | 25 | ABZ10100 | Haematopoietic cel |
| c | 22 | 90.6 | 5.6 | 8136 | 24 | ABK39957 | Human chemically p |
| c | 23 | 90.6 | 5.6 | 8136 | 24 | ABL32555 | Human immune syste |
| c | 24 | 90 | 5.6 | 11745 | 24 | ABK28332 | DNA transcription |
| c | 25 | 90 | 5.6 | 18060 | 24 | ABL92212 | Chemically treated |
| c | 26 | 90 | 5.6 | 18060 | 24 | AAD22313 | Chemically treated |
| c | 27 | 89.4 | 5.5 | 6621 | 24 | ABL70156 | Human DNA for stag |
| c | 28 | 89.4 | 5.5 | 6621 | 24 | ABK33933 | Human gene regulat |
| c | 29 | 89.4 | 5.5 | 6621 | 24 | AAS61098 | Human gene regulat |
| c | 30 | 89.4 | 5.5 | 6621 | 25 | ABZ09984 | Haematopoietic cel |
| c | 31 | 89.4 | 5.5 | 6621 | 25 | ABZ10098 | Haematopoietic cel |
| c | 32 | 89.4 | 5.5 | 6621 | 25 | ABZ10130 | Haematopoietic cel |
| c | 33 | 89.4 | 5.5 | 6621 | 25 | ABZ10244 | Haematopoietic cel |
| c | 34 | 89.4 | 5.5 | 113445 | 22 | AAI61373 | Soybean 318013 reg |
| c | 35 | 89.2 | 5.5 | 8136 | 24 | ABK39957 | Human chemically p |
| c | 36 | 89.2 | 5.5 | 8136 | 24 | ABL32555 | Human immune syste |
| c | 37 | 88.2 | 5.5 | 3296 | 23 | ABL17264 | Drosophila melanog |
| c | 38 | 87.4 | 5.4 | 6065 | 24 | ABL32505 | Human immune syste |
| c | 39 | 87.4 | 5.4 | 14551 | 24 | ABL34585 | Human metastasis a |
| c | 40 | 87.2 | 5.4 | 8136 | 24 | ABK39956 | Human chemically p |
| c | 41 | 87.2 | 5.4 | 8136 | 24 | ABL32554 | Human immune syste |
| c | 42 | 87.2 | 5.4 | 14551 | 24 | ABL34585 | Human metastasis a |
| c | 43 | 87.2 | 5.4 | 18855 | 24 | ABL32611 | Human immune syste |
| c | 44 | 86.8 | 5.4 | 169739 | 24 | ABQ86186 | Human osteoblast d |
| c | 45 | 86.2 | 5.3 | 20420 | 22 | AAK73165 | Human immune/haema |

ALIGNMENTS

RESULT 1
AAZ34539
ID AAZ34539 standard: DNA; 1612 bp.
XX AAZ34539;
AC AAZ34539;
XX 01-FEB-2000 (first entry)
XX Alfalfa salt inducible MsPRP2 gene promoter region.
DE Alfalfa salt inducible MsPRP2 gene promoter region.
XX MsPRP2 gene; promoter; Alfin1; transcription factor; alfalfa;
KW salt tolerance; stress tolerance; transgenic plant; root; ds.
XX Medicago sativa.
OS Medicago sativa.
FH Key Location/Qualifiers
FT protein_bind complement (718..727)
FT /*tag= a
FT /*note= "Alfin1 binding site"
FT complement (778..786)
FT /*tag= b
FT /*note= "Alfin1 binding site"
FT complement (1034..1049)
FT /*tag= c
FT /*note= "Alfin1 binding site"
FT complement (1079..1087)
FT /*tag= d
FT /*note= "Alfin1 binding site"
FT complement (1155..1160)
FT /*tag= e
FT /*note= "Alfin1 binding site"

Db 1342 CAATTGCGATATTATGATTTAGATTGCTGCAATACGGTCGCGTAATGTGATCACT 1401
 QY 1403 CACGAGAAAGAGGTATCAAAATTTCAAGGTATTTTATTTATTTTAAACAATAAATTTTC 1462
 Db 1402 CACGAGAAAGAGGTATCAAAATTTCAAGGTATTTTATTTTATTTTAAACAATAAATTTTC 1461
 QY 1463 AAGGTCCTTGTACCATATAAACCCTCCCTCACTCACACCCCAATTCCTTAAAGTGTATGACT 1522
 Db 1462 AAGGTCCTTGTACCATATAAACCCTCCCTCACTCACACCCCAATTCCTTAAAGTGTATGACT 1521
 QY 1523 TCATAGTACACTACACTACTTCTTTGAAACATGCTCACTATGCTCTAGCCAAATGTTTT 1582
 Db 1522 TCATAGTACACTACACTACTTCTTTGAAACATGCTCACTATGCTCTAGCCAAATGTTTT 1581
 QY 1583 CATCCTCTCTTGAACCTTGAGTACCTTACTC*1613
 Db 1582 CATCCTCTCTTGAACCTTGAGTACCTTACTC 1612

RESULT 2
 ID ABZ10246/c
 ID ABZ10246 standard; DNA; 8056 BP.
 XX AC ABZ10246;
 XX DT 16-JAN-2003 (first entry)
 XX DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX OS Homo sapiens.
 XX PN WO200277272-A2.
 XX PD 03-OCT-2002.
 XX PF 26-MAR-2002; 2002WO-EP03401.
 XX PR 26-MAR-2001; 2001US-278333P.
 XX PA (EPIG-) EPIGENOMICS AG.

XX PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
 PI Pelet C, Schwöbe I, Ziebarth H;
 XX WPI; 2003-018942/01.
 XX DR
 XX PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX PS Claim 28; SEQ ID 386; 117pp; English.
 XX CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder

CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX Sequence 8056 BP: 3711 A; 0 C; 371 G; 3974 T; 0 other;
 SQ Query Match 7.1%; Score 114; DB 25; Length 8056;
 Best Local Similarity 44.6%; Pred. No. 5.7e-07;
 Matches 613; Conservative 1; Mismatches 751; Indels 9; Gaps 4;
 QY 1 TTTTATAATATTTAAAGCTTGATTAATTTTGGCATCTATATATAAGCCCACTACCAAT 60
 Db 1855 TTTTATTTTATTTTAAACAAAATTTTATTTATTTAAATTAATAATAATA 1796
 QY 61 TTTAAAT 120
 Db 1795 ACATAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1736
 QY 121 TTTAAAT 180
 Db 1735 TTTAAAT 1679
 QY 181 TTACCTGTCATGCTAAAAAATAAATATATATTTTGTACCATTTTAAAGTCATATAATAGT 240
 Db 1678 TAAATTTTATCATATTAATTAATAATTTTATTTTAAATTTTACAAATTTTATTTCAA 1619
 QY 241 ACAATCCZACCCCTTTGAGAGGTTAATGTGTGCGGATTTTCTAGATAAACAGGCGCA 300
 Db 1618 TTTTTFATAATATTAATAATTAATAATTAATAATTAATAATTAATAATTAATTTT 1559
 QY 301 TTCACGATCTCTTGGTGACGCTTGGAGAACCCCTATCTCGGCTTGGAGATTTACTTC 360
 Db 1558 TAAATAAATAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1499
 QY 361 TTGTTGATGCTCTAGAGTACAGCTCTTAAAGCTGTAGTCTAGTCTAGTCTAGTCTCT 420
 Db 1498 ATTAATTTAATAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1439
 QY 421 TCCTACCAAAAAAAGTCAATAATAGTTTATATACATATAAATTTTAAATAAATAA 480
 Db 1438 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1379
 QY 481 AAAAATTCATCCCTAAAAACATAGTAAATTTTCAAAAAAATAATTTGTTTATATTT 540
 Db 1378 AATAAATTTTAAATAAATAAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1319
 QY 541 ACATCCGTTACGGTAAAAAATGGATAAATTTGGGTATGGGTATGAGTACTAGTAAATTAAGT 600
 Db 1318 TTTTATTAATTTTAAATTTTAAATAATATATACATTTTATATATTTAATAATAATACA 1259
 QY 601 TCATTGGTTAAAAAATCT 660
 Db 1258 TAACATTTTATTTTATTTCAAAAAATTTATTTTATTTTAAATAAATAAATAAATAA 1199
 QY 661 TTGGAACATGAAGGGTATTTGATTTTACACCTTTTACACCTTTTCAAGCCATTCAGA 720
 Db 1198 TTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1139
 QY 721 TCAATATATATTTTGGGGGATCAACACCAAGATCATTTACCATACATGCTTATACATA 780
 Db 1138 TCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1079
 QY 781 CCCCGTCATCTCTCTTTTACCAATTAACATTTTAAACATTTTAAACATTTTAAACAT 840
 Db 1078 TAAACAAATTTTATTTTATTTTACATTAATTAATTTCAAAATAAATAAATAAATAA 1019
 QY 841 ATAAAAACATCAAGCTAGCAAAATTTGTTTGGCATGACATTTTCATATAGTTTA 900

Db 1018 TTTTAAATTAATATTTTAAATAATATAAT-ATATATTTTATTCATTTAAATA 963
 QY 901 AAGGATGCATGATTCGATTACAAAACAAAATCTATATAATCTAGCACAAGGTTAAAG 960
 Db 962 AAAAATAAATATTTTAAATAAATTTTATACAAAACAAAATTAATTAATAA 903
 QY 961 CAATATATTAAGCTTCATAGCATGTGGATTTTCATTTAGAAAATATAGATTGCCCC 1020
 Db 902 TATAAATTTAAATAATTTACAAAATTTTATACATAAAAAAATAAAAAACAATTTT 843
 QY 1021 CTTTCATCAGGGTCTAACAGCACCCTTGTCTACTAGATGTCAAAATGTCCTCTAGTAC 1080
 Db 842 TTTAATAAATAATTTTAAATAATCATTTATAAATTTTATTTAAATATATAATTAATAAT 783
 QY 1081 AGCACCCTTTTACATTGATTCCTTGTCCATGCATGAATAAATAAATAAATAATTTG 1140
 Db 782 AATTTTATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 723
 QY 1141 GACACACAACTTGCCTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
 Db 722 TAAATTAATTTTAAATATATAAATAATTTTATATATTTTATTTTATTTTATTTT 663
 QY 1201 ATTGA-TCAAATTTGGCTATGAATTCACAAAATAATTCCTCTCTCTCTCTCTCTCT 1259
 Db 662 ATAAATTTATTTTAAATAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAAATA 603
 QY 1260 GGGCCACATATAAATCCATGAAGGATTTCAAT-GTCCATCCCAAGTCAATGATTCACAT 1318
 Db 602 TAAATATTTTAAATAATTAATAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAAAT 543
 QY 1319 ATATACATTTGAATAATTTAAATTTCCAAATTTGCAATTTATGATTTAGATTTGATT 1372
 Db 542 TTTAATTTTAAATAATTAATAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAAATTTT 489

RESULT 3
 ABL70376/c
 ID ABL70376 standard; DNA; 16258 BP.
 AC ABL70376;
 XX
 XX
 XX 01-JUL-2002 (first entry)
 XX
 XX Chemically treated cell signalling DNA sequence complementary to#133.
 XX
 XX Cell signalling; cytosine methylation; cell signalling disease;
 XX cancer; tumour; cytostatic; ds.
 XX
 XX Unidentified.
 XX
 XX W0200202807-A2.
 XX
 XX 10-JAN-2002.
 XX
 XX 29-JUN-2001; 2001WO-EP07471.
 XX
 XX 30-JUN-2000; 2000DE-1032529.
 XX
 XX 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 XX
 XX WPI; 2002-154758/20.
 XX
 XX Nucleic acid, useful for diagnosis and therapy of diseases associated
 XX with cell signalling e.g. cancer, comprises chemically modified genomic
 XX sequences of genes associated with cell signalling
 XX
 XX Claim 1; SEQ ID NO 266; 24ppr-sequence listing; English.
 XX
 XX The invention relates to a nucleic acid comprising a sequence of at least
 XX 18 bases of a segment of chemically pretreated DNA of genes associated

CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 XX Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;
 SQ
 Query Match 6.2%; Score 100.6; DB 24; Length 16258;
 Best Local Similarity 49.0%; Pred. No. 2.9e-05;
 Matches 381; Conservative 0; Mismatches 389; Indels 7; Gaps 4;
 QY 22 ATAAATTTTGGCATCTATATAGCCCACTACCAATTTAAATTTATATATATATA 81
 Db 2382 ATAAATATATAATATATATATTTCTACCTATATATAAATAAATAATATATATA 2323
 QY 82 TATATATATATATATATAATTAATTTTATTTTACCAATTTAAATTTATATATATA 141
 Db 2322 TA-ATATATATAAATAATATATAATATATATAATATATATATATATATA 2264
 QY 142 TATATATATATATATATAATTAATTTTATTTATTTTATTTAGTGGTGGTAAAAAAT 201
 Db 2263 TATACATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2204
 QY 202 AAATATAATTTGTTTACCATTTTAAAGTCATAAATATAGTACA-ATCAACCCCTTTGAGAG 260
 Db 2203 ATATATATCTAAACATATAAATAATATAAATAATATAATATATATATATATAA 2144
 QY 261 GTTAATGTGTGCGGATTTTCTAGATAAACAAGGCGCCATTCACGATTTCTTGGTGC 320
 Db 2143 ATATATACGTATATATATAATATATATATATATATATATATATATATATATAT 2084
 QY 321 AGCTTGGAGNACCCCTATCTCGGCTTGGAGATTTTACTTCTTGTGATCTCTAGA --- 377
 Db 2083 ATAAATATATACATATAAATAAACAAGTATATATATATATATATATATATATATA 2024
 QY 378 -GTACAGCTCTTAAGGCTGTAGTCTAGTGTGTTTCTCTCTCTCTCTCTCTCTCTCT 436
 Db 2023 TATATATAATATATATATACCATATAAATAATATATATATATATATATATATAT 1964
 QY 437 AAAGTCATAAATAGTTTATACATATACTTTTAAATAAATAAATAAATAAATAAATA 496
 Db 1963 TATATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1904
 QY 497 AAACATAGTAGAAATTTTCATAAATAAATAAATAAATAAATAAATAAATAAATAA 555
 Db 1903 TATAATATATATATATATATATATAAATAAATAAATAAATAAATAAATAAATAA 1844
 QY 556 AAAAAATGGATAAATTTGGGTATGGAGTACTAGTAAATAAATAAATAAATAAATAA 615
 Db 1843 TATATATATATATATATATATATATATAAATAAATAAATAAATAAATAAATAA 1784
 QY 616 ACTAAAAATAATTTCT 675
 Db 1783 AT 1724
 QY 676 TATTGATTTTACCACTTTTACACCTTTCAAGCCATTCACAGGATCAATATAGATTTT 735
 Db 1723 CCAAACTCTCTCAAAATCTACTCTCAAAATCTACTCTCAAACTCTCAAACTCTCA 1664
 QY 736 GGGCGATCAACACAAAGAATCATTAGGATAAATGCTTATACATACCCCTCAATCT 792
 Db 1663 ACAAACAAATAATATCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1607

RESULT 4
 ABR40038/C
 ID ABR40038 standard; DNA; 16258 BP.
 XX
 AC ABR40038;
 XX
 AC
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DT
 XX
 DT
 XX
 DE Human chemically pretreated gene sequence #60 strand 2.
 XX
 DE
 XX
 KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
 KW Cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
 KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
 XX
 OS Homo sapiens.
 XX
 OS
 XX
 PN W0200202806-A2.
 XX
 PN
 XX
 PD 10-JAN-2002.
 XX
 XX
 XX 29-JUN-2001; 2001WO-EP07470.
 XX
 XX
 XX 30-JUN-2000; 2000DE-1032529.
 XX
 XX 01-SEP-2000; 2000DE-1043826.
 XX
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 XX
 XX WPI; 2002-154757/20.
 XX
 XX
 XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
 XX useful for detecting cytosine methylation state of genes associated
 XX with pharmacogenomics and for therapy of diseases e.g. cancer -
 XX
 XX
 XX Cláim 1; SEQ ID No 120; 24pp; English.
 XX
 XX The invention relates to a nucleic acid comprising a sequence at
 XX least 18 bases in length of a segment of the chemically pretreated DNA
 XX of genes associated with pharmacogenomics according to one of the
 XX sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
 XX (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
 XX (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
 XX MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
 XX NM_019899) and their complementary sequences, or a sequence (SI) chosen
 XX from 87 sequences and their complements. The chemical pretreatment
 XX is bisulphite treatment to convert cytosines (but not methyl-cytosines)
 XX into uracils. Also included are an oligomer (II) in particular an
 XX oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
 XX each case at least one base sequence having a length of 9 nucleotides
 XX which hybridises to or is identical to a chemically pretreated DNA of
 XX genes associated with pharmacogenomics and their complements, arranged in
 XX an array for analysing diseases associated with the methylation state
 XX (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
 XX of the 87 sequences. The oligomers may also be used as PCR primers.
 XX The set of 87 nucleic acids and their complements is useful for diagnosis
 XX and therapy of solid tumours and cancer. The present sequence
 XX represents one the 87 DNA sequences or its complement.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 XX Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;
 XX

| | | | | | |
|----|-----------------------|--|--------------------|--------|---------------|
| | Query Match | 6.2% | Score 100.6; | DB 24; | Length 16258; |
| | Best Local Similarity | 49.0%; | Pred. No. 2.9e-05; | | |
| | Matches 381; | Conservative | 0; Mismatches 389; | Indels | 7; Gaps 4; |
| QY | 22 | ATAAATATTTGGCGATCTATATATAAGCCGCACATACCAAATTTAAAAATTTATATATATATATA 81 | | | |
| | | | | | |
| Dd | 2382 | TATATATATATATATATATATTTTCACCPATATATAAAAAAAAATAAATATATATATATA 2323 | | | |
| | | | | | |

| | | | |
|--------------|------|---|------|
| Qy | 82 | TATATATATATATATATATAAATAAAATTTTATTATTACCAATTTAAAAATATATATATATATA | 141 |
| Db | 2322 | TA-ATATATATAAAAATATATATATAAATATATATAAATATATATAAATATATATATATATA | 2264 |
| Qy | 142 | TATATATATATATATATATAAATAAATTTTTTATATATATTTATTAGCTTGATGGTAAAAAAT | 201 |
| Db | 2263 | TATACATATATATAAANAATAAATATATATATATATAAATATATATAAANAATAT | 2204 |
| Qy | 202 | AAATATAAATGTTGCCATTTTAAAGTCATAAATATAGTAGACA-ATCCAACCCCTTTGAGAG | 260 |
| Db | 2203 | ATATATATACTAAAAACATATAAAAAATATAAATAATATATAATATATATACATATAAAAAAC | 2144 |
| Qy | 261 | GTTAAATCTGTGCGGATTTTCTAGATAAACACAGGYGCCATTCACGAFTTCTTCTGGTGC | 320 |
| Db | 2143 | ATATATACGTATATATATAAATATATATATACATTTTAAAAACGTATATATATATAT | 2084 |
| Qy | 321 | AGCTTGGAGAACCTATCTCGGGCTTGGAGAATTTACTTCTTGTGATGCTTCAG- -- | 377 |
| Db | 2083 | ATAATATATACATATAAAAAACGATATATATATAAATTTATATACATATAAAAA | 2024 |
| Qy | 378 | -GTACAGCICCTTAAAGCTGTAGTCTAGTTTTTTTTTTTTCATCCTTCCTCACCAAAAAAAA | 436 |
| Db | 2023 | TATATATATAATATATATACCATAAAAAATATATATATTTATTATACATAAAAAACGTA | 1964 |
| Qy | 437 | AAAGTCATAAATATAGTTTTATACATATAAATCTTTTAATAAAAAATAAAAAATTTTCATCCCTA | 496 |
| Db | 1963 | TATATATATATATAAATAATATATATATATATATATATAAATAATATATATATATATATA | 1904 |
| Qy | 497 | AAAAACATATAGAAAAATTTCTCAAAAAAATAT-TGTTTATAAATTTACATGCCGTTACGGT | 555 |
| Db | 1903 | TATAAATATATATATATATATATAAATAATATATATATATATATATATATATATATAA | 1844 |
| Qy | 556 | AAAAATGATAAATGGGGPATGGAGTACGTAGTAATTAATAAAGTTCATTTGGTTAAAAAAA | 615 |
| Db | 1843 | TATATATATAAATATATATATATATATATATATAAATAATATATATATATATATATAAT | 1784 |
| Qy | 616 | ACTAAAAAATAATTTCTCTCCTGATTTTATATGAAATGCACATTTTTTTTGGAAACATGAAGG | 675 |
| Db | 1783 | ATATATATAAATAATTTTACTCTCA | 1724 |
| Qy | 676 | TATTGATTTTACCACCTTTTACACCTTTTCAAAGCCATTCGAAGGATGAATATAGATTTTT | 735 |
| Db | 1723 | CCAAAACCTCTTACAAATACTACTCTTCAAATAATTCACCTCCACCCATCTAAAAACCCAT | 1664 |
| Qy | 736 | GGGGGATCAACACAGAAATCATACGATAACATGCTTTATACATACCCCCGTCATCT | 792 |
| Db | 1663 | AACAAAACIAAATAATCATTTATTTCTCCCTTTCTTAACCTCCTCACTTACCTCTCT | 1607 |
| RESULT 5 | | | |
| ABZ10246 | ID | ABZ10246 standard; DNA; 8056 BP. | |
| XX | XX | ABZ10246; | |
| XX | XX | | |
| DT | DT | 16-JAN-2003 (first entry) | |
| XX | XX | | |
| DE | DE | Haematopoietic cell proliferation disorder related DNA sequence #386. | |
| XX | XX | | |
| KW | KW | Human; haematopoietic cell proliferation disorder; cytostatic; | |
| KW | KW | gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia; | |
| XX | XX | cytosine methylation state; gene; ds. | |
| OS | OS | Homo sapiens. | |
| XX | XX | | |
| PN | PN | WO200277272-A2. | |
| XX | XX | | |
| PD | PD | 03-OCT-2002. | |
| XX | XX | | |
| PF | PF | 26-MAR-2002; 2002WO-EP03401. | |
| XX | XX | | |
| PR | PR | 26-MAR-2001; 2001US-278333P. | |

DR WPI: 2003-018942/01.
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent
 PT that distinguishes between methylated and non-methylated CpG
 PT dinucleotides -
 XX
 PS Claim 28; SEQ ID 240; 117pp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related
 CC DNA sequences. The nucleotide sequences from the present invention can
 CC also be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables
 CC a highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients.
 XX
 SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 other;
 Query Match 5.9%; Score 95.2; DB 25; Length 8056;
 Best Local Similarity 44.4%; Id: Mismatches 679; Indels 9; Gaps 4;
 Matches 551; Conservative 1; Mismatches 679; Indels 9; Gaps 4;
 1 TTTTATAATATTAAGCTTCATATATATTTGGCATCTATATATAAGCCACCTACCAAT 60
 DB 1855 TTTTATTTTATTTTATTTTAAACGAAATTTTATTTATTTATTAATTAATATAATA 1796
 61 TTAAT 120
 DB 1795 ACGTAATTAATTTTATTTTAAATTTTAAATTTTATTTTATTTTATTTTAAATATA 1736
 121 TTAAT 180
 DB 1735 TTAAT 1679
 181 TTACGTTGATGGTAAAAAATAAATATATATTTGGTACCATTTAAAGTCATAAATATAGT 240
 DB 1678 TAAATTTTATCGTATTATTAATAATATTTTATTTTATTTTATTTTATTTTATTTT 1619
 241 ACAATCCACCCCTTGAGAGGTTAATGTGTGGCGGATTTCTAGATAAACAAGGCGCA 300
 DB 1618 TTTTAAATATATTAATAATATATATATATATATATATATATATATATATATATAT 1559
 301 TTCAGGATCTCTGTGGTGGCTTGGAGACCCCTATCCTGGCTTGGAGATTTTACTTC 360
 DB 1558 TAAATAAATAATTTTATTTTAAATTTTATTTTAAAGTATTTATATATATATATATAT 1499
 361 TTGTTGATGCTCTAGACATACAGCTCCCTTAAGGCTGTAGTCTAGTTTTTTTTTTCATCCT 420
 DB 1498 ATTAATTTAATAATTAATTTTAAATATATTTTATTTTAAATAAATATTTTATTTT 1439
 421 TCCTACCAAAAAAATAAAGTCATAAATATAGTTTATATATATATATATATATATATAT 480
 DB 1438 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1379
 481 AA-AAATTTTCCCTTAAAAACATAGTAGAATTTTCAATAAATAAATAAATAAATAAAT 539
 DB 1378 AATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1319

QY 540 TACATGCCTTACGGTAAAAAATGGATAAATTTGGGTATGGAGTACTAGTAATTAATAAGG 599
 DB 1318 TTTTATAAATTTATTAATTTTAAATAATATATAGTTTTTATATATTAATAATATACG 1259
 QY 600 TTCATTGTTAAAAAATAAATAAATTTCTCTCGATTTATATGAAATGACATTTT 659
 DB 1258 TAAGCTT-TTTTATTTTTCGAAAAATTAATTT-TATTAATAAATAAATAAATAAATAA 1200
 QY 660 TTTGGAACATGAAGGTATTGATTTTACCACCTTTTACACCTTTTCAAGCCATCAAGG 719
 DB 1199 ATTAATAATTTTAAATAAATAAATAATCGAAATATATTAATTTTAAATTTTCGAAT 1140
 QY 720 ATGAATAVAGATTTTGGCGCATCAACACAGAATCATTTACGATAAATGCTTATACAT 779
 DB 1139 ATCGAAAAAATAATTAATAACGATATATAACGTTTAAACGTTTAAAAAATAACGNA 1080
 QY 780 ACCCGGTCAATCTCTTTTATCCCAATTAACATTAAGATTTGCTTCTTTCGTTAAG 839
 DB 1079 TTAACACGAAATTTTATTTTACGTAATTAATTTTAAATTTTAAATAAATAAATAA 1020
 QY 840 CATAAAAACATCAAGTCTAGCAAAATGTTGTTTGGCATGACACATTTTCATATAGTT 899
 DB 1019 ATTTTAAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 964
 QY 900 AAAGGATCATGATTCGATTTACAAAAACAAAAATACTAATAATCTAGCACAAAGTTTAA 959
 DB 963 AAAAAAATAAATAATTTTATTTTAAATAAATAATTTATTCGAAAAACGCAATAATAA 904
 QY 960 GCAATATATTAAGCTTCATAGCATGTGATATTCATTTAGAAATATAGATAGATGCG 1019
 DB 903 ATATAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 844
 QY 1020 CCTTTCATCAGGGTCTAACAGCACCACTTGTCACTACATGTCACAAATGCTCTACTA 1079
 DB 843 TTTTAATAAATAAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 784
 QY 1080 CAGCACCGCTTTTACTTGTGATTCCTGTCATGCATGCAAAAAAATAAATAAATAA 1139
 DB 783 TAATTTTATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 724
 QY 1140 GGACACAACTTCCGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1199
 DB 723 TTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 664
 QY 1200 TATTGATCAATTTGGCTATGAATTCACAAACAAAAATTCA 1239
 DB 663 AATAAATTTAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 624
 RESULT 7
 ABL54336
 ID ABL54336 standard; DNA; 6641 BP.
 XX
 AC ABL54336;
 XX
 DT 29-JUL-2002 (first entry)
 XX
 DE Chemically treated apoptosis gene complementary to gene #18.
 XX
 KW Apoptosis; HIV; Bloom syndrome; cardiopathy;
 KW neurodegenerative disorder; Herpes simplex virus; renal ischaemia;
 KW amyotrophic lateral sclerosis; cancer; ds.
 XX
 OS Unidentified.
 XX
 PN WO200177164-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03969.
 XX
 PR 06-APR-2000; 2000HE-1019058.

PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPiG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-017444/02.
 XX
 XX Chemically modified sequences of genes associated with apoptosis are
 PT useful to determine methylation patterns of genomic DNA samples for
 PT diagnosis of associated diseases such as cancer.
 XX
 PS Claim 1; Seq ID #36; 24pp; English.
 XX
 CC This invention relates to chemically pre-treated DNA of genes
 CC associated with apoptosis. The nucleic acids are used to allocate
 CC patients for specific therapy for HIV infection, Bloom syndrome,
 CC cardiopathy, aging, neurodegenerative disorders, Herpes simplex virus
 CC infection, renal ischaemia, amyotrophic lateral sclerosis, solid tumours
 CC and cancers. This nucleotide sequence represents a chemically
 CC treated apoptosis gene. Even SEQ ID numbers are the complementary
 CC DNA strands to the odd SEQ ID numbers. The sequence data for this
 CC patent is not represented in the printed specification but is based on
 CC information supplied by the European patent office.
 XX
 SQ Sequence 6641 BP; 2495 A; 22 C; 984 G; 3140 T; 0 other;

Query Match 5.88; Score 93.2; DB 24; Length 6641;
 Best Local Similarity 46.58; Pred. No. 0.00031;
 Matches 451; Conservative 0; Mismatches 503; Indels 16; Gaps 4;
 QY 12 TTTAAGCTTCATATATATTTGGGCTATATATA-AGCCCACTACCAATTTAAATAT 70
 DB 2349 TATAATTTAGATTTTATTTAGGAGATATATATATATATATATATATATATAT 2408
 QY 71 AT 130
 DB 2409 AT 2468
 QY 131 AT 190
 DB 2469 AT 2525
 QY 191 GGTAAAAAATAATATATATTTGTTACCATTTAAAAAGTCATAAATAGTACAACTCAAC 250
 DB 2526 TAT 2585
 QY 251 CCTTTGAGAGTTAATGTGTGCGGATTTCTAGATAAACAAGGCCATTCACGATTC 310
 DB 2586 TAT 2645
 QY 311 TTTCTTGGTCGAGCTTGGAGACCCCTATCTGGCTGGAGATTTACTCTCTGTTGATGC 370
 DB 2646 TTAATAAAGAGATATCATTTAAGTAGTGTGTATGTTGGAAGCTTAGATATTTTT-T 2701
 QY 371 TTTCTAGAGTACACTCCTTAAGGCTCTAGTCTAGTTTTTTTTTTTTCATCCTCTACCAA 430
 DB 2702 TTTTGTGTAAATTTTATAAGATTTATTTAGTTTATTTAATATAAGGATAGTTTTT 2761
 QY 431 AAAAAAAGTCATAAATATAGTTTACATATATATATATATATATATATATATATATATAT 490
 DB 2762 ATATAAATATATTTTTTAAAGAAATA-----TGTAGGATATTCGAATAAATTTT 2813
 QY 491 TCCCTAAAAACATAGTAGAATTTTCATAAAAAAATATTTGTTTATATTTACATGCCGTT 550
 DB 2814 GTTTGAAGATATAGATTTATTTTTTTAGGTTTATTTTATATATATATATATATATAT 2873
 QY 551 ACGGTAAAAAAGGATAATTTGGGTATGGAGTACTAGTAATTAATAAGCTTCATGGTTA 610
 DB 2874 TAATTTATTTAGAAAGATATATGGAAGTTATAAATAGGATAAAGATATATATATATA 2933

QY 611 AAAAAACTAAAAAATATATTTCTCTCCGTGATTTATATGAAATGACATTTTTTTGGAAACATG 670
 DB 2934 AATGTAATATTTTTTTATGATTTTTTTTCTGATTAATAAATCAATTTATATAAAGTAATGAT 2993
 QY 671 AAGGTATTCATTTTACACCTTTTACACCTTTTCAAGCCNTTCAAGGATGATATATAGA 730
 DB 2994 AATTATATATATATATAGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3053
 QY 731 TTTTGGGCGCATCAACACACAAAGATCATTTACGATACATCTTATACATACCCCGTCAAT 790
 DB 3054 TTATATATGATTTAATTTAGTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 3113
 QY 791 CTCTCTTTTTTACCCTAATAACATTTGAAATGTTGCTTCTTCTGTTAAGCATAAAAACAT 850
 DB 3114 ATTATATATATGATATAAATAATAGTAATTTATTTATTTATTTATTTATTTATTTATTT 3173
 QY 851 CAAAGTCTAGCAAAATGTTGTTTTCGGATGACACATTTTCATATAGTTTTTAAAGATGCAT 910
 DB 3174 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3233
 QY 911 GATTGATTTACAAAAACAAATACTAATAATTTCTAGCACAAAGTTTAAAGCAATATTTATA 970
 DB 3234 TTAAGTATTTTATAAATAGTAGTAGTATTTGATTTTATTTATTTATTTATTTATTTAT 3293
 QY 971 AAGCTTCATA 980
 DB 3294 TATTTTATTA 3303

RESULT 8
 ABL32315
 ID ABL32315 standard; DNA; 6641 BP.
 XX
 AC ABL32315;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 288.
 DE
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; anti-naemic; cytosine; neurotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antichratic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200200928-A2.
 XX
 XX 03-JAN-2002.
 XX
 XX 02-JUL-2001; 2001WO-EP07537.
 XX
 XX 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation.
 PS
 PS Claim 1; SEQ ID NO 288; 32pp + Sequence Listing; German.
 XX
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences

CC amplicates carry a detectable label. The method further involves
 CC identifying methylation status of one or more cytosine positions, and
 CC analysing methylation status of the cytosine positions by reference to
 CC one or more data sets. The genomic DNA is chemically treated by using a
 CC bisulphite, hydrogen sulphite or disulphite. The amplification
 CC step amplifies DNA which is of particular interest in astrocytoma or
 CC brain tissue, based on the specific genomic methylation status of brain
 CC tissues, as opposed to background DNA. The amplicates carry a
 CC fluorescent label or radionuclide. Optionally, the labels of the
 CC amplicates are detachable molecule fragments having a typical mass
 CC which are detected in a mass spectrometer. The fragments of chemically
 CC pre-treated genomic DNA to be amplified, have a single positive or
 CC negative charge for a better detectability in the mass spectrometer.
 CC Preferably, the amplicates or fragments of the amplicates are
 CC detected by matrix assisted laser desorption/ionization mass spectrometry
 CC (MALDI) or using electron spray mass spectrometry (ESI). The
 CC present sequence is one of the chemically pre-treated reference DNA
 CC samples of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 6621 BP; 1770 A; 214 C; 1781 G; 2856 T; 0 other;

Query Match 5.7%; Score 92; DB 24; Length 6621;
 Best Local Similarity 61.9%; Pred. No. 0.00044;
 Matches 146; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2 TTTTAAATTTTAAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTCAATTT 61
 Db 4889 TGTGTGATATATAATAATATATATATATGTGTATATATAATAATATATATATG 4948
 QY 62 TAAATTTAT 121
 Db 4949 TGTA 5008
 QY 122 TAAATTTAT 181
 Db 5009 TAT 5068
 QY 182 TACGTTGATGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA 237
 Db 5069 GTAAT 5124

RESULT 12

AAS61098
 ID AAS61098 standard; DNA; 6621 BP.

XX AC AAS61098;

XX DT 29-JAN-2002 (first entry)

XX DE Human gene regulation-associated gene oligonucleotide #53.

KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX OS Homo sapiens.

XX PN WO20017375-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-EP03968.

XX PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-017470/02.

XX New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease -

PS Claim 1; SEQ ID No 54; 26pp; English.

XX The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 6621 BP; 1770 A; 214 C; 1781 G; 2856 T; 0 other;

Query Match 5.7%; Score 92; DB 24; Length 6621;
 Best Local Similarity 61.9%; Pred. No. 0.00044;
 Matches 146; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2 TTTTAAATTTTAAAGCTTGATATAATTTTGGCATCTATATATAAGCCCACTCAATTT 61
 Db 4889 TGTGTGATATATAATAATATATATATATGTGTATATATAATAATATATATATG 4948
 QY 62 TAAATTTAT 121
 Db 4949 TGTA 5008
 QY 122 TAAATTTAT 181
 Db 5009 TAT 5068
 QY 182 TACGTTGATGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA 237
 Db 5069 GTAAT 5124

RESULT 13

ABZ09984

ID ABZ09984 standard; DNA; 6621 BP.

XX AC ABZ09984;

XX DT 16-JAN-2003 (first entry)

XX Haematopoietic cell proliferation disorder related DNA sequence #124.

XX Human; haematopoietic cell proliferation disorder; cytostatic;

THIS PAGE BLANK (USPTO)


```

; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801.861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match
Best Local Similarity 4.7%; Score 75.6; DB 4; Length 53332;
Matches 138; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACCAATT 61
DB 31616 TATATAAATATATAAATATATAAATATAAATATAAATATAAATATAAATATAAATAA 31557

QY 62 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 31556 TATATAAATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAT 31497

QY 122 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 181
DB 31496 AAATATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAAT 31437

QY 182 TACGTTGATGTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACCA 241
DB 31436 ATAAATATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATA 31377

QY 242 CA 243
DB 31376 TA 31375

```

```

RESULT 3
US-09-783-203-1/c
; Sequence 1, Application US/09783203
; Patent No. 6576464
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gold, Joseph
; APPLICANT: Lebkowski, Jane
; TITLE OF INVENTION: TPACKED stem cells
; FILE REFERENCE: 096/003
; CURRENT APPLICATION NUMBER: US/09/783.203
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/253,443
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-783-203-1

Query Match
Best Local Similarity 4.7%; Score 75.4; DB 4; Length 15418;
Matches 136; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 2 TTTATAAATATTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACCA 61
DB 1185 TTTATATATGATATATTTGATATATATATATATATATATATATATATATATATATAT 1126

QY 62 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 1125 TTTATATATGATATATTTTATATATATATATATATATATATATATATATATATATAT 1066

QY 122 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 181

```

```

DB 1065 TATATATGATATATTTTATATATATATATATATATATATATATATATATATATATAT 1006
QY 182 TACGTTGATGTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACCA 238
DB 1005 TATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT 949

RESULT 4
US-09-783-203-1
; Sequence 1, Application US/09783203
; Patent No. 6576464
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gold, Joseph
; APPLICANT: Lebkowski, Jane
; TITLE OF INVENTION: TPACKED stem cells
; FILE REFERENCE: 096/003
; CURRENT APPLICATION NUMBER: US/09/783.203
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/253,443
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,357
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 15418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-783-203-1

```

```

Query Match
Best Local Similarity 4.6%; Score 73.8; DB 4; Length 15418;
Matches 157; Conservative 0; Mismatches 102; Indels 2; Gaps 2;

QY 2 TTTATAAATATTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACCA 61
DB 979 TATATACATATATAAATATATATTTTACATATATAAATATATACATATAAATA-TACA 1037

QY 62 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 121
DB 1038 TAAATACATATATAAATATATACATATATAAATATATACATATAAATATACATATAA 1097

QY 122 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 180
DB 1098 TATACATATATAAATATATACATATATAAATATATATATATATATATATATATATA 1157

QY 181 TTACGTTGATGTTAAAGCTTGATAATAATTTTGGCGATCTATATATAAGCCCACTACA 240
DB 1158 AGTATATACAAATATATACATATATAAATGATATATAGTATATACATATATATAAATA 1217

QY 241 ACAATCCACCCCTTTGAGAGG 261
DB 1218 TATAAAAAAACTTTTGGCTGG 1238

```

```

RESULT 5
US-09-671-317-141
; Sequence 141, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLIS
; FILE REFERENCE: 62.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/671.317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536.178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403

```

APPLICANT: Su, Xin-zhaun

ORGANISM: Plasmodium falciparum

[illegible]

102 ; NUMBER OF SEQ ID NOS: 134
103 ; SOFTWARE: Patent.pm
104 ; SEQ ID NO 1
105 ; LENGTH: 319608
106 ; TYPE: DNA
107 ; ORGANISM: Homo sapiens
108 ; FEATURE:
109 ; NAME/KEY: misc_feature
110 ; LOCATION: 199122..201122
111 ; OTHER INFORMATION: 5'regulatory region
112 ; NAME/KEY: exon
113 ; LOCATION: 201123..201234
114 ; OTHER INFORMATION: exon S
115 ; NAME/KEY: exon
116 ; LOCATION: 201123..201560
117 ; OTHER INFORMATION: exon S2
118 ; NAME/KEY: exon
119 ; LOCATION: 214676..214793
120 ; OTHER INFORMATION: exon T
121 ; NAME/KEY: exon
122 ; LOCATION: 215702..215746
123 ; OTHER INFORMATION: exon U
124 ; NAME/KEY: exon
125 ; LOCATION: 216836..216994
126 ; OTHER INFORMATION: exon V
127 ; NAME/KEY: exon
128 ; LOCATION: 216836..217077
129 ; OTHER INFORMATION: exon V2
130 ; NAME/KEY: exon
131 ; LOCATION: 217671..217764
132 ; OTHER INFORMATION: exon V1
133 ; NAME/KEY: exon
134 ; LOCATION: 227655..227736
135 ; OTHER INFORMATION: exon V4
136 ; NAME/KEY: exon
137 ; LOCATION: 238715..238919
138 ; OTHER INFORMATION: exon V3
139 ; NAME/KEY: exon
140 ; LOCATION: 240440..240673
141 ; OTHER INFORMATION: exon W
142 ; NAME/KEY: exon
143 ; LOCATION: 240440..241153
144 ; OTHER INFORMATION: exon W2
145 ; NAME/KEY: exon
146 ; LOCATION: 241072..241291
147 ; OTHER INFORMATION: exon X
148 ; NAME/KEY: exon
149 ; LOCATION: 244353..244561
150 ; OTHER INFORMATION: exon Y
151 ; NAME/KEY: exon
152 ; LOCATION: 246273..247802
153 ; OTHER INFORMATION: exon Z
154 ; NAME/KEY: misc_feature
155 ; LOCATION: 247803..249803
156 ; OTHER INFORMATION: 3'regulatory region
157 ; NAME/KEY: allele
158 ; LOCATION: 8316
159 ; OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
160 ; NAME/KEY: allele
161 ; LOCATION: 21672
162 ; OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
163 ; NAME/KEY: allele
164 ; LOCATION: 65485
165 ; OTHER INFORMATION: 8-128-33 : polymorphic base C or T
166 ; NAME/KEY: allele
167 ; LOCATION: 95396
168 ; OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
169 ; NAME/KEY: allele
170 ; LOCATION: 107281
171 ; OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
172 ; NAME/KEY: allele
173 ; LOCATION: 160640
174 ; OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
175 ;
176 ;
177 ;
178 ;
179 ;
180 ;
181 ;
182 ;
183 ;
184 ;
185 ;
186 ;
187 ;
188 ;
189 ;
190 ;
191 ;
192 ;
193 ;
194 ;
195 ;
196 ;
197 ;
198 ;
199 ;
200 ;
201 ;
202 ;
203 ;
204 ;
205 ;
206 ;
207 ;
208 ;
209 ;
210 ;
211 ;
212 ;
213 ;
214 ;
215 ;
216 ;
217 ;
218 ;
219 ;
220 ;
221 ;
222 ;
223 ;
224 ;
225 ;
226 ;
227 ;
228 ;
229 ;
230 ;
231 ;
232 ;
233 ;
234 ;
235 ;
236 ;
237 ;
238 ;
239 ;
240 ;
241 ;
242 ;
243 ;
244 ;
245 ;
246 ;
247 ;
248 ;
249 ;
250 ;
251 ;
252 ;
253 ;
254 ;
255 ;
256 ;
257 ;
258 ;
259 ;
260 ;
261 ;
262 ;
263 ;
264 ;
265 ;
266 ;
267 ;
268 ;
269 ;
270 ;
271 ;
272 ;
273 ;
274 ;
275 ;
276 ;
277 ;
278 ;
279 ;
280 ;
281 ;
282 ;
283 ;
284 ;
285 ;
286 ;
287 ;
288 ;
289 ;
290 ;
291 ;
292 ;
293 ;
294 ;
295 ;
296 ;
297 ;
298 ;
299 ;
300 ;
301 ;
302 ;
303 ;
304 ;
305 ;
306 ;
307 ;
308 ;
309 ;
310 ;
311 ;
312 ;
313 ;
314 ;
315 ;
316 ;
317 ;
318 ;
319 ;
320 ;
321 ;
322 ;
323 ;
324 ;
325 ;
326 ;
327 ;
328 ;
329 ;
330 ;
331 ;
332 ;
333 ;
334 ;
335 ;
336 ;
337 ;
338 ;
339 ;
340 ;
341 ;
342 ;
343 ;
344 ;
345 ;
346 ;
347 ;
348 ;
349 ;
350 ;
351 ;
352 ;
353 ;
354 ;
355 ;
356 ;
357 ;
358 ;
359 ;
360 ;
361 ;
362 ;
363 ;
364 ;
365 ;
366 ;
367 ;
368 ;
369 ;
370 ;
371 ;
372 ;
373 ;
374 ;
375 ;
376 ;
377 ;
378 ;
379 ;
380 ;
381 ;
382 ;
383 ;
384 ;
385 ;
386 ;
387 ;
388 ;
389 ;
390 ;
391 ;
392 ;
393 ;
394 ;
395 ;
396 ;
397 ;
398 ;
399 ;
400 ;
401 ;
402 ;
403 ;
404 ;
405 ;
406 ;
407 ;
408 ;
409 ;
410 ;
411 ;
412 ;
413 ;
414 ;
415 ;
416 ;
417 ;
418 ;
419 ;
420 ;
421 ;
422 ;
423 ;
424 ;
425 ;
426 ;
427 ;
428 ;
429 ;
430 ;
431 ;
432 ;
433 ;
434 ;
435 ;
436 ;
437 ;
438 ;
439 ;
440 ;
441 ;
442 ;
443 ;
444 ;
445 ;
446 ;
447 ;
448 ;
449 ;
450 ;
451 ;
452 ;
453 ;
454 ;
455 ;
456 ;
457 ;
458 ;
459 ;
460 ;
461 ;
462 ;
463 ;
464 ;
465 ;
466 ;
467 ;
468 ;
469 ;
470 ;
471 ;
472 ;
473 ;
474 ;
475 ;
476 ;
477 ;
478 ;
479 ;
480 ;
481 ;
482 ;
483 ;
484 ;
485 ;
486 ;
487 ;
488 ;
489 ;
490 ;
491 ;
492 ;
493 ;
494 ;
495 ;
496 ;
497 ;
498 ;
499 ;
500 ;
501 ;
502 ;
503 ;
504 ;
505 ;
506 ;
507 ;
508 ;
509 ;
510 ;
511 ;
512 ;
513 ;
514 ;
515 ;
516 ;
517 ;
518 ;
519 ;
520 ;
521 ;
522 ;
523 ;
524 ;
525 ;
526 ;
527 ;
528 ;
529 ;
530 ;
531 ;
532 ;
533 ;
534 ;
535 ;
536 ;
537 ;
538 ;
539 ;
540 ;
541 ;
542 ;
543 ;
544 ;
545 ;
546 ;
547 ;
548 ;
549 ;
550 ;
551 ;
552 ;
553 ;
554 ;
555 ;
556 ;
557 ;
558 ;
559 ;
560 ;
561 ;
562 ;
563 ;
564 ;
565 ;
566 ;
567 ;
568 ;
569 ;
570 ;
571 ;
572 ;
573 ;
574 ;
575 ;
576 ;
577 ;
578 ;
579 ;
580 ;
581 ;
582 ;
583 ;
584 ;
585 ;
586 ;
587 ;
588 ;
589 ;
590 ;
591 ;
592 ;
593 ;
594 ;
595 ;
596 ;
597 ;
598 ;
599 ;
600 ;
601 ;
602 ;
603 ;
604 ;
605 ;
606 ;
607 ;
608 ;
609 ;
610 ;
611 ;
612 ;
613 ;
614 ;
615 ;
616 ;
617 ;
618 ;
619 ;
620 ;
621 ;
622 ;
623 ;
624 ;
625 ;
626 ;
627 ;
628 ;
629 ;
630 ;
631 ;
632 ;
633 ;
634 ;
635 ;
636 ;
637 ;
638 ;
639 ;
640 ;
641 ;
642 ;
643 ;
644 ;
645 ;
646 ;
647 ;
648 ;
649 ;
650 ;
651 ;
652 ;
653 ;
654 ;
655 ;
656 ;
657 ;
658 ;
659 ;
660 ;
661 ;
662 ;
663 ;
664 ;
665 ;
666 ;
667 ;
668 ;
669 ;
670 ;
671 ;
672 ;
673 ;
674 ;
675 ;
676 ;
677 ;
678 ;
679 ;
680 ;
681 ;
682 ;
683 ;
684 ;
685 ;
686 ;
687 ;
688 ;
689 ;
690 ;
691 ;
692 ;
693 ;
694 ;
695 ;
696 ;
697 ;
698 ;
699 ;
700 ;
701 ;
702 ;
703 ;
704 ;
705 ;
706 ;
707 ;
708 ;
709 ;
710 ;
711 ;
712 ;
713 ;
714 ;
715 ;
716 ;
717 ;
718 ;
719 ;
720 ;
721 ;
722 ;
723 ;
724 ;
725 ;
726 ;
727 ;
728 ;
729 ;
730 ;
731 ;
732 ;
733 ;
734 ;
735 ;
736 ;
737 ;
738 ;
739 ;
740 ;
741 ;
742 ;
743 ;
744 ;
745 ;
746 ;
747 ;
748 ;
749 ;
750 ;
751 ;
752 ;
753 ;
754 ;
755 ;
756 ;
757 ;
758 ;
759 ;
760 ;
761 ;
762 ;
763 ;
764 ;
765 ;
766 ;
767 ;
768 ;
769 ;
770 ;
771 ;
772 ;
773 ;
774 ;
775 ;
776 ;
777 ;
778 ;
779 ;
780 ;
781 ;
782 ;
783 ;
784 ;
785 ;
786 ;
787 ;
788 ;
789 ;
790 ;
791 ;
792 ;
793 ;
794 ;
795 ;
796 ;
797 ;
798 ;
799 ;
800 ;
801 ;
802 ;
803 ;
804 ;
805 ;
806 ;
807 ;
808 ;
809 ;
810 ;
811 ;
812 ;
813 ;
814 ;
815 ;
816 ;
817 ;
818 ;
819 ;
820 ;
821 ;
822 ;
823 ;
824 ;
825 ;
826 ;
827 ;
828 ;
829 ;
830 ;
831 ;
832 ;
833 ;
834 ;
835 ;
836 ;
837 ;
838 ;
839 ;
840 ;
841 ;
842 ;
843 ;
844 ;
845 ;
846 ;
847 ;
848 ;
849 ;
850 ;
851 ;
852 ;
853 ;
854 ;
855 ;
856 ;
857 ;
858 ;
859 ;
860 ;
861 ;
862 ;
863 ;
864 ;
865 ;
866 ;
867 ;
868 ;
869 ;
870 ;
871 ;
872 ;
873 ;
874 ;
875 ;
876 ;
877 ;
878 ;
879 ;
880 ;
881 ;
882 ;
883 ;
884 ;
885 ;
886 ;
887 ;
888 ;
889 ;
890 ;
891 ;
892 ;
893 ;
894 ;
895 ;
896 ;
897 ;
898 ;
899 ;
900 ;
901 ;
902 ;
903 ;
904 ;
905 ;
906 ;
907 ;
908 ;
909 ;
910 ;
911 ;
912 ;
913 ;
914 ;
915 ;
916 ;
917 ;
918 ;
919 ;
920 ;
921 ;
922 ;
923 ;
924 ;
925 ;
926 ;
927 ;
928 ;
929 ;
930 ;
931 ;
932 ;
933 ;
934 ;
935 ;
936 ;
937 ;
938 ;
939 ;
940 ;
941 ;
942 ;
943 ;
944 ;
945 ;
946 ;
947 ;
948 ;
949 ;
950 ;
951 ;
952 ;
953 ;
954 ;
955 ;
956 ;
957 ;
958 ;
959 ;
960 ;
961 ;
962 ;
963 ;
964 ;
965 ;
966 ;
967 ;
968 ;
969 ;
970 ;
971 ;
972 ;
973 ;
974 ;
975 ;
976 ;
977 ;
978 ;
979 ;
980 ;
981 ;
982 ;
983 ;
984 ;
985 ;
986 ;
987 ;
988 ;
989 ;
990 ;
991 ;
992 ;
993 ;
994 ;
995 ;
996 ;
997 ;
998 ;
999 ;
1000 ;
1001 ;
1002 ;
1003 ;
1004 ;
1005 ;
1006 ;
1007 ;
1008 ;
1009 ;
1010 ;
1011 ;
1012 ;
1013 ;
1014 ;
1015 ;
1016 ;
1017 ;
1018 ;
1019 ;
1020 ;
1021 ;
1022 ;
1023 ;
1024 ;
1025 ;
1026 ;
1027 ;
1028 ;
1029 ;
1030 ;
1031 ;
1032 ;
1033 ;
1034 ;
1035 ;
1036 ;
1037 ;
1038 ;
1039 ;
1040 ;
1041 ;
1042 ;
1043 ;
1044 ;
1045 ;
1046 ;
1047 ;
1048 ;
1049 ;
1050 ;
1051 ;
1052 ;
1053 ;
1054 ;
1055 ;
1056 ;
1057 ;
1058 ;
1059 ;
1060 ;
1061 ;
1062 ;
1063 ;
1064 ;
1065 ;
1066 ;
1067 ;
1068 ;
1069 ;
1070 ;
1071 ;
1072 ;
1073 ;
1074 ;
1075 ;
1076 ;
1077 ;
1078 ;
1079 ;
1080 ;
1081 ;
1082 ;
1083 ;
1084 ;
1085 ;
1086 ;
1087 ;
1088 ;
1089 ;
1090 ;
1091 ;
1092 ;
1093 ;
1094 ;
1095 ;
1096 ;
1097 ;
1098 ;
1099 ;
1100 ;
1101 ;
1102 ;
1103 ;
1104 ;
1105 ;
1106 ;
1107 ;
1108 ;
1109 ;
1110 ;
1111 ;
1112 ;
1113 ;
1114 ;
1115 ;
1116 ;
1117 ;
1118 ;
1119 ;
1120 ;
1121 ;
1122 ;
1123 ;
1124 ;
1125 ;
1126 ;
1127 ;
1128 ;
1129 ;
1130 ;
1131 ;
1132 ;
1133 ;
1134 ;
1135 ;
1136 ;
1137 ;
1138 ;
1139 ;
1140 ;
1141 ;
1142 ;
1143 ;
1144 ;
1145 ;
1146 ;
1147 ;
1148 ;
1149 ;
1150 ;
1151 ;
1152 ;
1153 ;
1154 ;
1155 ;
1156 ;
1157 ;
1158 ;
1159 ;
1160 ;
1161 ;
1162 ;
1163 ;
1164 ;
1165 ;
1166 ;
1167 ;
1168 ;
1169 ;
1170 ;
1171 ;
1172 ;
1173 ;
1174 ;
1175 ;
1176 ;
1177 ;
1178 ;
1179 ;
1180 ;
1181 ;
1182 ;
1183 ;
1184 ;
1185 ;
1186 ;
1187 ;
1188 ;
1189 ;
1190 ;
1191 ;
1192 ;
1193 ;
1194 ;
1195 ;
1196 ;
1197 ;
1198 ;
1199 ;
1200 ;
1201 ;
1202 ;
1203 ;
1204 ;
1205 ;
1206 ;
1207 ;
1208 ;
1209 ;
1210 ;
1211 ;
1212 ;
1213 ;
1214 ;
1215 ;
1216 ;
1217 ;
1218 ;
1219 ;
1220 ;
1221 ;
1222 ;
1223 ;
1224 ;
1225 ;
1226 ;
1227 ;
1228 ;
1229 ;
1230 ;
1231 ;
1232 ;
1233 ;
1234 ;
1235 ;
1236 ;
1237 ;
1238 ;
1239 ;
1240 ;
1241 ;
1242 ;
1243 ;
1244 ;
1245 ;
1246 ;
1247 ;
1248 ;
1249 ;
1250 ;
1251 ;
1252 ;
1253 ;
1254 ;
1255 ;
1256 ;
1257 ;
1258 ;
1259 ;
1260 ;
1261 ;
1262 ;
1263 ;
1264 ;
1265 ;
1266 ;
1267 ;
1268 ;
1269 ;
1270 ;
1271 ;
1272 ;
1273 ;
1274 ;
1275 ;
1276 ;
1277 ;
1278 ;
1279 ;
1280 ;
1281 ;
1282 ;
1283 ;
1284 ;
1285 ;
1286 ;
1287 ;
1288 ;
1289 ;
1290 ;
1291 ;
1292 ;
1293 ;
1294 ;
1295 ;
1296 ;
1297 ;
1298 ;
1299 ;
1300 ;
1301 ;
1302 ;
1303 ;
1304 ;
1305 ;
1306 ;
1307 ;
1308 ;
1309 ;
1310 ;
1311 ;
1312 ;
1313 ;
1314 ;
1315 ;
1316 ;
1317 ;
1318 ;
1319 ;
1320 ;
1321 ;
1322 ;
1323 ;
1324 ;
1325 ;
1326 ;
1327 ;
1328 ;
1329 ;
1330 ;
1331 ;
1332 ;
1333 ;
1334 ;
1335 ;
1336 ;
1337 ;
1338 ;
1339 ;
1340 ;
1341 ;
1342 ;
1343 ;
1344 ;
1345 ;
1346 ;
1347 ;
1348 ;
1349 ;
1350 ;
1351 ;
1352 ;
1353 ;
1354 ;
1355 ;
1356 ;
1357 ;
1358 ;
1359 ;
1360 ;
1361 ;
1362 ;
1363 ;
1364 ;
1365 ;
1366 ;
1367 ;
1368 ;
1369 ;
1370 ;
1371 ;
1372 ;
1373 ;
1374 ;
1375 ;
1376 ;
1377 ;
1378 ;
1379 ;
1380 ;
1381 ;
1382 ;
1383 ;
1384 ;
1385 ;
1386 ;
1387 ;
1388 ;
1389 ;
1390 ;
1391 ;
1392 ;
1393 ;
1394 ;
1395 ;
1396 ;
1397 ;
1398 ;
1399 ;
1400 ;
1401 ;
1402 ;
1403 ;
1404 ;
1405 ;
1406 ;
1407 ;
1408 ;
1409 ;
1410 ;
1411 ;
1412 ;
1413 ;
1414 ;
1415 ;
1416 ;
1417 ;
1418 ;
1419 ;
1420 ;
1421 ;
1422 ;
1423 ;
1424 ;
1425 ;
1426 ;
1427 ;
1428 ;
1429 ;
1430 ;
1431 ;
1432 ;
1433 ;
1434 ;
1435 ;
1436 ;
1437 ;
1438 ;
1439 ;
1440 ;
1441 ;
1442 ;
1443 ;
1444 ;
1445 ;
1446 ;
1447 ;
1448 ;
1449 ;
1450 ;
1451 ;
1452 ;
1453 ;
1454 ;
1455 ;
1456 ;
1457 ;
1458 ;
1459 ;
1460 ;
1461 ;
1462 ;
1463 ;
1464 ;
1465 ;
1466 ;
1467 ;
1468 ;
1469 ;
1470 ;
1471 ;
1472 ;
1473 ;
1474 ;
1475 ;
1476 ;
1477 ;
1478 ;
1479 ;
1480 ;
1481 ;
1482 ;
1483 ;
1484 ;
1485 ;
1486 ;
1487 ;
1488 ;
1489 ;
1490 ;
1491 ;
1492 ;
1493 ;
1494 ;
1495 ;
1496 ;
1497 ;
1498 ;
1499 ;
1500 ;
1501 ;
1502 ;
1503 ;
1504 ;
1505 ;
1506 ;
1507 ;
1508 ;
1509 ;
1510 ;
1511 ;
1512 ;
1513 ;
1514 ;
1515 ;
1516 ;
1517 ;
1518 ;
1519 ;
1520 ;
1521 ;
1522 ;
1523 ;
1524 ;
1525 ;
1526 ;
1527 ;
1528 ;
1529 ;
1530 ;
1531 ;
1532 ;
1533 ;
1534 ;
1535 ;
1536 ;
1537 ;
1538 ;
1539 ;
1540 ;
1541 ;
1542 ;
1543 ;
1544 ;
1545 ;
1546 ;
1547 ;
1548 ;
1549 ;
1550 ;
1551 ;
1552 ;
1553 ;
1554 ;
1555 ;
1556 ;
1557 ;
1558 ;
1559 ;
1560 ;
1561 ;
1562 ;
1563 ;
1564 ;
1565 ;
1566 ;
1567 ;
1568 ;
1569 ;
1570 ;
1571 ;
1572 ;
1573 ;
1574 ;
1575 ;
1576 ;
1577 ;
1578 ;
1579 ;
1580 ;
1581 ;
1582 ;
1583 ;
1584 ;
1585 ;
1586 ;
1587 ;
1588 ;
1589 ;
1590 ;
1591 ;
1592 ;
1593 ;
1594 ;
1595 ;
1596 ;
1597 ;
1598 ;
1599 ;
1600 ;
1601 ;
1602 ;
1603 ;
1604 ;
1605 ;
1606 ;
1607 ;
1608 ;
1609 ;
1610 ;
1611 ;
1612 ;
1613 ;
1614 ;
1615 ;
1616 ;
1617 ;
1618 ;
1619 ;
1620 ;
1621 ;
1622 ;
1623 ;
1624 ;
1625 ;
1626 ;
1627 ;
1628 ;
1629 ;
1630 ;
1631 ;
1632 ;
1633 ;
1634 ;
1635 ;
1636 ;
1637 ;
1638 ;
1639 ;
1640 ;
1641 ;
1642 ;
1643 ;
1644 ;
1645 ;
1646 ;
1647 ;
1648 ;
1649 ;
1650 ;
1651 ;
1652 ;
1653 ;
1654 ;
1655 ;
1656 ;
1657 ;
1658 ;
1659 ;
1660 ;
1661 ;
1662 ;
1663 ;
1664 ;
1665 ;
1666 ;

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|--------|--------|-------------------|----|--------------------|
| | | Match | % | | | | |
| 1 | 93.2 | 5.8 | 6641 | 12 | US-10-311-455-288 | | Sequence 288, App |
| 2 | 92.4 | 5.7 | 9770 | 12 | US-10-311-455-5 | | Sequence 5, Appli |
| 3 | 92 | 5.7 | 6621 | 14 | US-10-172-086-10 | | Sequence 10, Appl |
| 4 | 91.6 | 5.7 | 11745 | 12 | US-10-240-453-206 | | Sequence 206, App |
| 5 | 91.4 | 5.7 | 6641 | 12 | US-10-311-455-288 | | Sequence 288, App |
| 6 | 91 | 5.6 | 3991 | 14 | US-10-074-045-60 | | Sequence 60, Appl |
| 7 | 90.6 | 5.6 | 8136 | 12 | US-10-311-455-528 | | Sequence 528, App |
| 8 | 90 | 5.6 | 11745 | 12 | US-10-240-453-206 | | Sequence 206, App |
| 9 | 89.4 | 5.5 | 6621 | 14 | US-10-172-086-10 | | Sequence 10, Appli |
| 10 | 88.4 | 5.5 | 513509 | 11 | US-09-754-855A-4 | | Sequence 4, Appli |
| 11 | 88.2 | 5.5 | 8136 | 12 | US-10-311-455-528 | | Sequence 528, App |
| 12 | 87.4 | 5.4 | 6065 | 12 | US-10-311-455-478 | | Sequence 478, App |
| 13 | 87.4 | 5.4 | 14551 | 12 | US-10-240-485-138 | | Sequence 138, App |
| 14 | 87.2 | 5.4 | 8136 | 12 | US-10-311-455-527 | | Sequence 527, App |
| 15 | 87.2 | 5.4 | 14551 | 12 | US-10-240-485-138 | | Sequence 138, App |
| 16 | 87.2 | 5.4 | 18855 | 12 | US-10-311-455-584 | | Sequence 584, App |

2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500


```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT221C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 3991
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-045-60

```

[illegible]

RESULT 7
 US-10-311-455-528/c
 ; Sequence 528, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
 ; FILE OF INVENTION: cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT APPLICATION NUMBER: US/10/311,455
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 528
 ; LENGTH: 8136
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-311-455-528

| | Query Match | 5.6% | Score 90.6 | DB 12 | Length 8136 |
|----|-----------------------|---|-------------------|----------|-------------|
| | Best Local Similarity | 61.8% | Pred. No. 0.00065 | | |
| | Matches 144 | Conservative 0 | Mismatches 89 | Indels 0 | Gaps 0 |
| QY | 15 | AAGCTTGATAAATAATTTTCGGATCTATATATATAGCCCACTACCAATTTAAAAATTATATAT | 74 | | |
| | | | | | |
| Db | 7316 | ACCTTAATACAAATAAACACTCAATCTCAAAAAAANAANAANAANAATATATAT | | | |
| | | | | | |
| QY | 75 | ATATATATATATATATATATATATAATAATTTTTTTATTTATCCCAATTTAAAAATTATATAT | 134 | | |
| | | | | | |

[illegible]

RESULT 8
US-10-240-453-206
: Sequence 206, Application US/10240453
: Publication No. US20030148326A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
: TITLE OF INVENTION: Transcription
: TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
: TITLE OF INVENTION: With DNA Transcription
: FILE REFERENCE: 5013.1009
: CURRENT APPLICATION NUMBER: US/10/240,453
: CURRENT FILING DATE: 2002-10-02
: PRIOR APPLICATION NUMBER: PCT/EP01/03973
: PRIOR FILING DATE: 2001-04-06
: PRIOR APPLICATION NUMBER: DE 10019058.8
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: DE 10019173.8
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: DE 10032529.7
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: DE 10043826.1
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 350

TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (9105)
 US-10-240-453-206

| Query Match | 5.6% | Score 90 | DB 12 | Length 11745 |
|-----------------------|----------------|---|----------|--------------|
| Best Local Similarity | 68.0% | Pred. No. 0.00087 | | |
| Matches 140 | Conservative 0 | Mismatches 65 | Indels 1 | Caps 1 |
| Qy | 4 | TATAAATATTTAAGCTTGATGAATAAATTTTCGGATCTATATATAAGGCCACTACCAATT | 63 | |
| Db | 8520 | TTTAGTTTTTAAAGGATTTTTTTCGAATTAATATATATATATATATATATATATATATA | 8579 | |
| Qy | 64 | AAATTTATATATATATATATATATATATATATATATATATATAAATTTTTATTTACCAATT | 123 | |
| Db | 8580 | TA | 8639 | |
| Qy | 124 | AAATTTATATATATATATATATATATATATATATATATAAATTTTTAT-TATATTTATT | 182 | |
| Db | 8640 | TA | 8699 | |
| Qy | 183 | ACGTTGATGGTAAAAAATAAATATA | 208 | |
| Db | 8700 | TATATATATATATATATATATATATA | 8725 | |

RESULT 9
US-10-172-086-10/c
; Sequence 10, Application US/10172086
; Publication No. US20030113750A1

[illegible]

```

RESULT 12
US-10-311-455-478
; Sequence 478, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 478
; LENGTH: 6065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-478

```

[illegible][illegible]

```

RESULT 13
US-10-240-485-138
; Sequence 138, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 138
; LENGTH: 14551
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo
US-10-240-485-138

```

| Query Match | 5.4% | Score 87.4 | DB 12 | Length 14551 |
|-----------------------|----------------|--|----------|--------------|
| Best Local Similarity | 64.7% | Pred. No. 0.0021 | | |
| Matches 130 | Conservative 0 | Mismatches 71 | Indels 0 | Gaps 0 |
| Qy | 15 | AAGCTTGATAAATATTTGGGATCTATATATAAGCCACATACCAATTTAAAAATTATATAT | 74 | |
| Db | 2750 | AATTATAATATTGATTAAGTTTTTATTTAGTTATTAGTTGGTATATATATATATAT | 2809 | |
| Qy | 75 | ATATATATATATATATATATATATAATATTTTATTTACCACATTTAAAAATTATATAT | 134 | |
| Db | 2810 | AT | 2869 | |
| Qy | 135 | ATATATATATATATATATATATATAAATATTTTATTTATTTATTTACGTTGATGGTA | 194 | |
| Db | 2870 | ATATATATATATATATATATATATATATATATATTTTGAATTTAATTATAAGTTAGTGTT | 2929 | |
| Qy | 195 | AAAAAATAAATAAATTTGTT | 215 | |
| Db | 2930 | ATAAAGAAGCAGTTTATGAT | 2950 | |

362 TGTGATGCTTCTAGAGTACAGCTCCTTAAGGCTGTAGTCTAGTTTTCATCCTT 421

RESULT 14

```
; S-10-311-455-527/c
; Sequence 527, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 527
; LENGTH: 8136
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-527
```

```
Query Match          5.4%; Score 87.2; DB 12; Length 8136;
Best Local Similarity 70.7%; Pred. No. 0.002;
Matches 116; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 23 TAATAAATTTTCGATCATATATAAGCCACACCAATTTAAATTTATATATATATATATATAT 82
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1014 TCATAAATTTACATTTTATATTTAAATCTATATATATATATATATATATATATACACAC 955

QY 83 ATATATATATATATATATAAATTTTATTTTACCAATTTAAATTTATATATATATATATAT 142
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 954 ACGTATATATATATATATATATATATATATATATATATACACATATATATATATATACACATAT 895

QY 143 ATATATATATATATATAAATTTTATTTATTTATTTATTTATTTACGT 186
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 894 ATATATATATATATATATATATTTTATTTTATTTTATTTTATTTTCT 851
```

RESULT 15

```
; US-10-240-485-138/c
; Sequence 138, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240.485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 138
; LENGTH: 14551
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

; FEATURE:

```
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-138

Query Match          5.4%; Score 87.2; DB 12; Length 14551;
Best Local Similarity 66.5%; Pred. No. 0.0023;
Matches 125; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 58 AATTAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 117
   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
DB 2910 AATTCAATATATATATATATATATATATATATATATATATATATATATATATATATATA 2851

QY 118 AATTAAATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 177
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
DB 2850 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 2791

QY 178 TTATTACGTTTCATGCTAAATAAATAAATAAATTTGTTACCATTTTAAAGTCATAAATAT 237
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
DB 2790 CTAATAACCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2731

QY 238 AGTACAAT 245
   |||||
DB 2730 ACAAAATAT 2723

Search completed: August 14, 2003, 23:27:34
Job time : 439 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 16:08:55 ; Search time 3661 seconds

(without alignments)
10708.309 Million cell updates/sec

Title: US-09-647-841B-1

Perfect score: 1613

Sequence: 1 ttttataatatttaagctt.....tgaacttgagctacttactc 1613

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:
1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: em_gss_hum:**
18: em_gss_inv:**
19: em_gss_pln:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_mam:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rod:**
26: em_gss_phg:**
27: em_gss_vrl:**
28: gb_gss1:**
29: gb_gss2:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 118.8 | 7.4 | 1201 | 9 | AL565455 |
| 2 | 115 | 7.1 | 1201 | 9 | AL565455 |
| 3 | 109.6 | 6.8 | 1200 | 13 | BX437758 |
| 4 | 105.8 | 6.6 | 1200 | 13 | BX415878 |

| | | | | | |
|----|-------|-----|------|----|-----------|
| 5 | 105.8 | 6.6 | 1200 | 13 | BX415878 |
| 6 | 104.6 | 6.5 | 1101 | 29 | CNS000EVL |
| 7 | 103 | 6.4 | 400 | 29 | DR46120T |
| 8 | 102.8 | 6.4 | 1056 | 13 | BX415058 |
| 9 | 102 | 6.3 | 797 | 28 | AQ868440 |
| 10 | 101.6 | 6.3 | 490 | 29 | BX237235 |
| 11 | 100.6 | 6.2 | 581 | 29 | BX217227 |
| 12 | 100.6 | 6.2 | 707 | 29 | BX153480 |
| 13 | 100 | 6.2 | 698 | 29 | BX200889 |
| 14 | 100 | 6.2 | 1201 | 13 | BX446511 |
| 15 | 99 | 6.1 | 605 | 29 | BX162842 |
| 16 | 98.4 | 6.1 | 1101 | 29 | CNS000EVL |
| 17 | 98.2 | 6.1 | 652 | 29 | BX217370 |
| 18 | 98.2 | 6.1 | 739 | 29 | BZ202464 |
| 19 | 98 | 6.1 | 797 | 28 | AQ868440 |
| 20 | 98 | 6.1 | 1200 | 13 | BX437758 |
| 21 | 97.6 | 6.1 | 371 | 29 | BX241244 |
| 22 | 97.4 | 6.0 | 488 | 29 | BX145735 |
| 23 | 97.2 | 6.0 | 675 | 29 | DR48P23T |
| 24 | 96.8 | 6.0 | 558 | 29 | DR39M22T |
| 25 | 96.8 | 6.0 | 574 | 29 | BX149536 |
| 26 | 96.8 | 6.0 | 605 | 29 | BX162842 |
| 27 | 96.6 | 6.0 | 548 | 29 | BX185514 |
| 28 | 96.4 | 6.0 | 772 | 29 | BX205878 |
| 29 | 96.2 | 6.0 | 400 | 29 | DR46120T |
| 30 | 96.2 | 6.0 | 604 | 29 | BX204248 |
| 31 | 96.2 | 6.0 | 781 | 29 | CC134240 |
| 32 | 96.2 | 6.0 | 829 | 29 | BX143241 |
| 33 | 96 | 6.0 | 697 | 29 | CNS04707 |
| 34 | 95.8 | 5.9 | 499 | 29 | BX162335 |
| 35 | 95.6 | 5.9 | 697 | 29 | CNS04707 |
| 36 | 95.4 | 5.9 | 404 | 29 | BX176026 |
| 37 | 95.4 | 5.9 | 478 | 29 | DR1G4T |
| 38 | 95.4 | 5.9 | 520 | 29 | BX224356 |
| 39 | 95.4 | 5.9 | 656 | 29 | DR36P11T |
| 40 | 95.4 | 5.9 | 739 | 29 | BZ202464 |
| 41 | 95.2 | 5.9 | 486 | 29 | BX225691 |
| 42 | 95.2 | 5.9 | 631 | 28 | AZ133337 |
| 43 | 95.2 | 5.9 | 658 | 29 | BX184360 |
| 44 | 95 | 5.9 | 594 | 29 | BX139459 |
| 45 | 95 | 5.9 | 648 | 29 | DR61B17T |

ALIGNMENTS

RESULT 1
AL565455/c
LOCUS
DEFINITION
AL565455 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION
AL565455
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On Feb 16, 2001 this sequence version replaced gi:12916848.
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: seqef@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9232.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF005H09NP1&cluster=9232.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDF005BH09NP1.
FEATURES

| | |
|--------|---|
| source | /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSDF005YO18" /tissue_type="FETAL BRAIN" /dev_stage="fetal" /clone_lib="Homo sapiens FETAL BRAIN" |
| | /note=Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized." |

```
BASE COUNT      550 a   82 C    71 G    313 T    185 others
ORIGIN
Query Match               7.4%; Score 118.8; DB 9; Length 1201;
Best Local Similarity     36.6%; Pred. No. 0.0052;
Matches 305; Conservative 136; Mismatches 388; Indels 5; Gaps 3.
```

| | | | |
|----|--------|---|------|
| QY | 54 | TACCAATTTAAAATTATATATATATATATATATATATATAAATTTTTATT | 113 |
| Dd | 1198 | TATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 1135 |
| QY | 114 | TACCAGTTTAAAAATATATATATATATATATATATATATATATATATATAT | 173 |
| Dd | 1138 | TATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 1135 |
| QY | 174 | ATATTTTATAGTGCTGGTAGAAATATAATATATATATATATATATATATAT | 233 |
| Dd | * 1079 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 1020 |
| QY | 234 | ATATAGTACAATCCAACCTTTGAGAGGTAACTGCTGCCGGATTTCTCAGATAAACAA | 293 |
| Dd | 1019 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 960 |
| QY | 294 | GCGCCACTCACGATCTCTCTGGTGGAGCGCTGGAGACCCCTACCTCGGCTGGAGCAT | 353 |
| Dd | 959 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 900 |
| QY | 354 | TTACTCTCTGGTGAAGCTTCAGAGTACAGCTCCCTTAAGGCTGAGCTAGTCTTTT | 413 |
| Dd | 899 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 840 |
| QY | 414 | TCATCTCTCTCACCAAATAAAGATGATGATGATGATGATGATGATGATGATGAT | 473 |
| Dd | 839 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 780 |
| QY | 474 | AAAATAAAAAAAAAATTCATCCCATAAAGATGATGATGATGATGATGATGATGAT | 533 |
| Dd | 779 | WTATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 721 |
| QY | 534 | ATAATTTACATGCGCTTACGGTAAAAAATGGATGATGATGATGATGATGATGAT | 593 |
| Dd | 720 | TWATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 661 |
| QY | 594 | ATAAGGTTCATTTGGTTAAAAAATATAATTTCTCCCTGCATTTATATGAAATGA | 653 |
| Dd | 660 | TWATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 601 |
| QY | 654 | CATTCTTTTGGACATGCAAGGTATGATTTTACCACCTTTTACACCTTTTCAAGGCAT | 713 |
| Dd | 600 | TWATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 544 |
| QY | 714 | TCAGGAGTGAATPAGATTTTGGCGCATCAAAACACAGATCATTCAGATAAAGCAT | 773 |
| Dd | 543 | ATATWTATWTAATGTATGATGCTGGAGCGCCTTCCTCCCTGCAGCATGAAGA | 484 |
| QY | 774 | ATACATACCCCGTCAATCTCTTTTACCACCAATAAAGATGATGATGATGATGAT | 833 |

[illegible]

[illegible]


```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP004AD10NP1.
Location/Qualifiers
1. .1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP004YGI9"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT-6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT      454 a   53 c   56 g   318 t   175 others
ORIGIN
Query Match      6.4%;   Score 102.8;   DB 13;   Length 1056;

```

[illegible][illegible][illegible]

| |
|---|
| AQ868440 |
| LOCUS |
| DEFINITION |
| AQ868440 linear DNA 797 bp GSS 03-NOV-1999 |
| nbe50027F06r CUGI Rice BAC Library (EcoRI) Oryza sativa [japonica |
| cultivar-group] genomic clone nbe50027F06r, genomic survey |
| sequence. |

ACCESSION AQ868440
 VERSION AQ868440.1 GI:6218891
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 797)
 REFERENCE Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished
 CONTACT: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence start: 41
 High quality sequence stop: 329.
 FEATURES
 source
 1..797
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbeb0027F06r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library (EcoRI)"
 /note="vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;
 Rice is the most important food crop in the world. Half of
 the world population, especially those inhabiting highly
 populated areas of the humid tropics and subtropics, rely
 on rice as their primary source of carbohydrate.
 Monocytodendous rice is a diploid plant (2n=24) with a
 haploid genome equivalent of 431 Mbp (Arumuganathan and
 Earle, 1991). The relatively small genome of rice, three
 times larger than that of Arabidopsis, makes it suitable
 for genomic studies. In order to facilitate positional
 cloning, physical mapping and genome sequencing of rice,
 we have constructed a BAC library from Oryza sativa,
 Nipponbare variety using EcoRI as the cloning enzyme. The
 library contains 55,296 clones with an average insert size
 of 121 kb providing approximately 15 haploid genome
 equivalents. The deep coverage allows the isolation a
 particular sequence with a probability of 99.9 %. Three
 high density filters, each containing 18,432 clones
 (doubly spotted), represent the whole library for colony
 screening and can be requested from the Clemson University
 BAC/EST Resource Center (www.genome.clemson.edu)."
 BASE COUNT 285 a 179 c 48 g 275 t 10 others
 ORIGIN
 Query Match 6.3%; Score 102; DB 28; Length 797;
 Best Local Similarity 59.8%; Pred. No. 0.25; Indels 0; Gaps 0;
 Matches 171; Conservative 0; Mismatches 115;
 Qy 2 TTTATAATATTTAAAGCTTGATATAATTTTGGCATCTATATATAATGACCCACTACCAATT 61
 Db 220 TGTAAATATATTTGTTTCATATATATATATCTATATACATATGCTTGGCATACATATA 279
 Qy 62 TAAATTTAT 121
 Db 280 TATGTATAAATACATA 339
 Qy 122 TAAATTTAT 181
 Db 340 TAT 399

182 TACCTTGATGGTAAAAAATAAATAATATATTTTGTACCAATTTTAAAGTCATAAATAGTA 241
 Db 400 AT 459
 Qy 242 CAATCCCAACCCCTTTGAGAGGTTAATGTGTGCGGATTTTCTAGAT 287
 Db 460 CACACACACAT 505
 RESULT 10
 BX237235 490 bp DNA linear GSS 29-JAN-2003
 LOCUS BX237235
 DEFINITION Danio rerio genomic clone DKEY-283H15, genomic survey sequence.
 ACCESSION BX237235
 VERSION BX237235.1 GI:28159569
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 490)
 REFERENCE Humphray, S.J., Huckle, E. and Durham, J.L.
 Direct SubMISSION
 AUTHORS
 TITLE
 JOURNAL
 COMMENT This sequence was generated from the SP6 end of BAC 283H15. 283H15
 is part of the Daniokey BAC Library created by R. Plasterk and N.V.
 Keygene. Further details:
 http://www.sanger.ac.uk/Projects/D_rerio/
 FEATURES
 source
 1..490
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-283H15"
 /tissue_type="Testis"
 /note="vector pIndigoBAC-536"
 BASE COUNT 204 a 32 c 35 g 219 t
 ORIGIN
 Query Match 6.3%; Score 101.6; DB 29; Length 490;
 Best Local Similarity 72.8%; Pred. No. 0.36;
 Matches 131; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Qy 59 ATTTAAAT 118
 Db 220 AT 279
 Qy 119 ATTTAAAT 178
 Db 280 AT 339
 Qy 179 TATTACCGTTGATGCTAAAAAATAAATAATATATTTTGTACCAATTTTAAAGTCATAAATATA 238
 Db 340 TATTTTA 399
 RESULT 11
 BX217227/c 581 bp DNA linear GSS 29-JAN-2003
 LOCUS BX217227
 DEFINITION Danio rerio genomic clone DKEY-258I22, genomic survey sequence.
 ACCESSION BX217227
 VERSION BX217227.1 GI:28049113
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 581)

| | |
|----------------|--|
| AUTHORS | Humphray, S. J., Huckle, E. and Durham, J. I. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrys@sanger.ac.uk Unpublished |
| COMMENT | This sequence was generated from the SP6 end of BAC 258I22. 258I22 is part of the Daniokey BAC Library created by R. Piasterk and N.V. keygene. Further details: http://www.sanger.ac.uk/projects/D_rerio/ . |

```

BASE COUNT
ORIGIN
208 a 60 c 81 g 232 t
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="pKEY-258122"
/tissue_type="Testis"
/notes="vector pindigORAC-536"

```

| | Query Match | 6.2% | Score 100.6; | DB 29; | Length 581; |
|----|-----------------------|--|-----------------|----------------|-------------|
| | Best Local Similarity | 65.5% | Pred. No. 0.42; | | |
| | Matches 163; | Conservative | 0; | Mismatches 84; | Indels 2; |
| | | | | Gaps | 1; |
| QY | 1 | TTTTATAAATATTATTAAGCTTGTAAATAATATTTTCGGATCTATATATATAGGCCCACTACCAAT | 60 | | |
| Db | 275 | TTGTGTGAATTTCTGATCACAGCACATATTTTGTAAACACATGTCTCTCTGTATAAT | 216 | | |

[illegible]

| | | | | | |
|------------|--------------------------|-------------|--------------------------|-----------------|--|
| RESULT 12 | | | | | |
| BX153480/c | | | | | |
| LOCUS | 707 bp | DNA | linear | GSS 13-MAR-2003 | |
| DEFINITION | DX10 rerio genomic clone | DKEY-114K5, | genomic survey sequence. | | |
| ACCESSION | DX153480 | | | | |

| | |
|-----------|---|
| KEYWORDS | GSS. |
| SOURCE | Danio rerio (zebrafish) |
| ORGANISM | Danio rerio |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. |
| REFERENCE | 1. (bases 1 to 707) |
| AUTHORS | Humphray, S.J., Huckle, E. and Durham, J.L. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished |
| COMMENT | This sequence was generated from the SP6 end of BAC 114K5. 114K5 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/projects/D_rerio/ . |
| FEATURES | Location/Qualifiers |

```
/organism="Danio rerio"  
/mol_type="genomic DNA"  
/db_xref="taxon:7955"  
/clone="DKEY-114K5"
```


THIS PAGE BLANK (USPTO)